iFW

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/038,977	12/31/2001	Douglas Neal Fuller	DF01-001	9586

7590 03/15/2005

Dr. Douglas Neal Fuller
P.O. Box 450936
Atlanta, GA 31145-0936

EXAMINER	
ZHU, JERRY	
ART UNIT	**PAPER NUMBER**
2121	

DATE MAILED: 03/15/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

RECEIVED
MAR 25 2005
Technology Center 2100

Office Action Summary	Application No. 10/038,977	Applicant(s) FULLER, DOUGLAS NEAL	
	Examiner Jerry Zhu	Art Unit 2121	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-21* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-21* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

the invention as disclosed in claims 1-21 is directed to non-statutory subject matter.

1. Claims 1-21 are method claims whose steps are not practiced on a computer, electronic devices, electrical machines, mechanical apparatus, or anything concrete and tangible instruments or equipments. These steps are just abstract procedures manipulating abstract concepts. Therefore, it is clear that these claims are not limited to practice in the technological arts. On that basis alone, they are clearly nonstatutory.
2. Regardless of whether any of the claims are in the technological arts, claims 1-21 are just manipulating abstract ideas. Congress intended statutory subject matter to 'include anything under the Sun that is made by man.'" ***Diamond v. Diehr,*** 450 U.S. at 182, 209 USPQ at 6. "This Court has undoubtably recognized limits to §101 and every discovery is not embraced within the statutory terms. Excluded from such patent protection are laws of

nature, physical phenomena and abstract ideas." ***Id.*** at 185, 209 USPQ at 7. A claim that covers any and every possible way that the steps can be performed is a disembodied "abstract idea" because there is no particular implementation of the idea. See Gottschalk vs. Benson, 409 U.S. at 68, 175 USPQ at 675 (The Supreme Court discussed the cases holding that a principle, in the abstract, cannot be patented and then stated: "Here is the 'process' claim is so abstract and sweeping as to cover both known and unknown uses of the BCD to pure binary conversion. The end use may ... be performed through any existing machinery or future-devised machinery or without any apparatus.")

Furthermore, in the case *In re Warmerdam*, the Federal Circuit held that:

> ...[T]he dispositive issue for assessing compliance with Section 101 in this case is whether the claim is for a process that goes beyond simply manipulating 'abstract ideas' or 'natural phenomena' ... As the Supreme Court has made clear, '[a]n idea of itself is not patentable, ... *taking several abstract ideas and manipulating them together adds nothing to the basic equation*. In re Warmerdam 31 USPQ2d at 1759 (emphasis added).

Since the Federal Circuit held in *Warmerdam* that this is the "dispositive issue" when it judged the usefulness, concreteness, and tangibility of the claim limitations in that case. Examiner in the present case views this holding as the dispositive issue for determining whether a claim is "useful, concrete, and tangible" in similar cases. Accordingly, the Examiner finds that the method claims manipulate a set of abstract ideas such as "population," "members," "rules," and "behavior." (i.e., what population it is? Population of marbles, animals, vehicles, people how have pets?) Clearly, manipulation of abstract ideas such as parameters, characteristics of abstract population is provably even more abstract (and thereby less limited in practical application) than pure "mathematical algorithms" which the Supreme Court has held are per se nonstatutory - in fact, it *includes* the expression of nonstatutory mathematical algorithms. Since the claims are not limited to exclude such abstractions, the broadest reasonable interpretation of the claim limitations includes such abstractions. Therefore, the claims are impermissibly abstract under 35 U.S.C. § 101.

3. Regardless of whether any of the claims are abstract nor not, none of them is limited to practical applications in the technological arts. There is no physical transformation either inside or outside of a computer as the result of performing the method. Examiner finds that *In re Warmerdam*, 33 F.3d 1354, 31 USPQ2d 1754 (Fed. Cir. 1994) controls the 35 USC §101 issues on that

point for reasons made clear by the Federal Circuit in *AT&T Corp. v. Excel Communications, Inc.*, 50 USPQ2d 1447 (Fed. Cir. 1999). Specifically, the Federal Circuit held that the act of:

> ...[T]aking several abstract ideas and manipulating them together adds nothing to the basic equation. *AT&T v. Excel* at 1453 quoting *In re Warmerdam*, 33 F.3d 1354, 1360 (Fed. Cir. 1994).

Examiner finds no evidence in the claims that manipulating "sub-population" using "rules" produces any concrete, tangible, practical, chemical, physical, or business transformation.

Examiner bases his position upon guidance provided by the Federal Circuit in *In re Warmerdam*, as interpreted by *AT&T v. Excel*. This set of precedents is within the same line of cases as the *Alappat-State Street Bank* decisions and is in complete agreement with those decisions. *Warmerdam* is consistent with *State Street*'s holding that:

> Today we hold that *the transformation of data, representing* _discrete dollar amounts_, *by a machine through a series of mathematical calculations into a final share price*, constitutes a practical application of a mathematical algorithm, formula, or calculation because it produces 'a useful, concrete and tangible result" -- *a final share price momentarily fixed for recording purposes and even accepted and relied upon by regulatory authorities and in subsequent trades*. (emphasis added) *State Street Bank* at 1601.

That case later eliminated the "business method exception" in order to show that business methods were not per se nonstatutory, but the court clearly *did not* go so far as to make business methods *per se statutory*. A plain reading of the excerpt above shows that the Court was *very specific* in its definition of the new *practical application*. It would have been much easier for the court to say that "business methods were per se statutory" than it was to define the practical application in the case as "...the transformation of data, representing discrete dollar amounts, by a machine through a series of mathematical calculations into a final share price..."

Additionally, the court was also careful to specify that the "useful, concrete and tangible result" it found was "a final share price momentarily fixed for recording purposes and even accepted and <u>relied upon</u> by regulatory authorities and in subsequent <u>trades</u>." (i.e. the trading activity is the <u>further practical use</u> of the real world <u>monetary</u> data beyond the transformation in the computer - i.e., "post-processing activity".)

Applicant cites no such specific results to define a useful, concrete and tangible result. Neither does Applicant specify the associated practical application with the kind of specificity the Federal Circuit used.

Assuming that the claims fall within the category of a "process" under §101, the steps are so broadly recited, without regard to any tangible way of implementing them, that they are directed to the "abstract idea" itself and the claims are nonstatutory subject matter under the "abstract idea" exception. The abstract ideas comprising the steps are not instantiated into some specific physical implementation. Nor are there any minor physical acts, such as recording, that might be construed as an implementation of the abstract idea.

Where a claim is broad enough to read on both statutory subject matter (machine implementation or physical transformation of physical subject matter) as well as nonstatutory subject matter (an abstract idea), the best position is to hold the claimed subject matter to be nonstatutory because, while a claim is a pending and can be amended, a claim's meaning should be delimited by express terms rather than claim interpretation. ***Cf. In re Lintner***, 458 F. 2d 1013, 1015, 173 USPQ 560, 562 (CCPA 1972) ("Claims which are broad enough to read on obvious subject matter are unpatentable even though they also read on non-obvious subject matter.").

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jerry Zhu whose telephone number is (571) 2724237. The examiner can normally be reached on 8:30 - 5.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Anthony Knight can be reached on (571) 272-3687. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Jerry Zhu
Examiner
Art Unit - 2121
2/1/2005

Anthony Knight
Supervisory Patent Examiner
Tech Center 2100

#2

PTO/SB/08A (10-01)
Approved for use through 10/31/2002. OMB 0651-0031
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 1 of 3

Complete if Known	
Application Number	
Filing Date	
First Named Inventor	FULLER
Art Unit	
Examiner Name	
Attorney Docket Number	DF01-001

10/038977 12/31/01

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number: Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
JZ		US- 6,272,483	8/7/2001	Joslin, Clements	
JZ		US- 6,269,339	7/31/2001	Silver	
JZ		US- 6,223,164	5/24/2001	Seare et al	
JZ		US- 6,088,510	7/11/2000	Sims	
JZ		US- 5,999,893	12/7/1999	Lynch et al	
JZ		US- 6,026,397	2/15/2000	Sheppard	
JZ		US- 6,272,478	8/7/2001	Obata et al	
JZ		US- 6,212,526	4/3/2001	Chaudhuri et al.	
JZ		US- 6,125,362	9/26/2000	Elworthy	
JZ		US- 5,787,420	7/28/2001	Tukey	
JZ		US- 5,712,984	1/27/1998	Hammond et al.	
JZ		US- 6,029,138	2/22/2000	Khorasami et al.	
JZ		US- 6,202,053	3/13/2001	Christiansen et al.	
JZ		US- 6,266,656	7/24/2001	Ohno	
JZ		US- 6,182,070	1/30/2001	Megiddo et al.	
JZ		US- 5,187,673	2/16/1993	Carver, Jr. et al.	
JZ		US- 5,790,758	8/4/1998	Streit	
		US-			
		US-			
		US-			

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document: Country Code3 -Number[4] - Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T[6]

Examiner Signature	[signature]	Date Considered	2/08/05

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

PTO/SB/08B (10-01)
Approved for use through 10/31/2002. OMB 0651-0031
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 2 of 3

Complete if Known	
Application Number	
Filing Date	
First Named Inventor	FULLER
Group Art Unit	
Examiner Name	
Attorney Docket Number	DF01-001

C932 U.S. PTO
12/31/01

OTHER PRIOR ART -- NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published	T[2]
JZ	1.	Anderberg, Michael, Cluster Analysis for Applications, Academic Press, New York, 1973.	
JZ	2.	Bock, HH, Classification and Related Methods of Data Analysis, Elsevier Science Publishers, North Holland, 1988	
JZ	3.	Celeux, Gilles and Soromenho, Gilda, "An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model," *Journal of Classification*, 13, 195-212, 1996	
JZ	4.	Fowles JB et al, "Taking health status into account when setting capitation rates: a comparison of risk-adjustment methods," *JAMA* 276(16): 1316-21, 1996.	
JZ	5.	Hornbrook, Mark C & Goodman Michael J., "Assessing Relative Health Plan Risk with the Rand 36 Health Survey," *Inquiry*, 32:1, 56-74, 1995.	
JZ	6.	Hornbrook, Mark C & Goodman Michael J., "Chronic Disease, Functional Health Status, and Demographics: A Multi-Dimensional Approach to Risk Adjustment," *Health Services Research*, 31:3, 283-307, 1996.	
JZ	7.	Lamers, LM & van Vliet RC, "Multiyear diagnostic information from prior hospitalization as a risk-adjuster for capitation payments," *Medical Care*, 34:4 549-561, 1996.	
JZ	8.	Lorr, Maurice, Cluster Analysis for Social Scientists, Jossey-Bass Publishers, San Francisco,1983.	
JZ	9.	Matthews, Georffrey & Hearne, James, "Clustering Without a Metric," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, V13N2, p175-184, 1991.	
JZ	10.	Soromenho, Gilda, "Comparing Approaches for Testing the Number of Components in a Finite Mixture Model," *Computational Statistics*, 9, 65-78,1994.	
JZ	11.	Nord, Erik, "Health Status Index Models for use in Resource Allocation Decisions," *International Journal of Technology Assessment in Health Care*, 12:1, 1996.	

Examiner Signature	[signature]	Date Considered	2/28/05

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] Applicant is to place a check mark here if English language Translation is attached.

PTO/SB/08B (10-01)
Approved for use through 10/31/2002. OMB 0651-0031
U.S. Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

Sheet 3 of 3

Complete if Known	
Application Number	
Filing Date	
First Named Inventor	FULLER
Group Art Unit	
Examiner Name	
Attorney Docket Number	DF01-001

10/028977 U.S. PTO
12/31/01

OTHER PRIOR ART – NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published	T[2]
JZ	12.	Anderson, Richard V., "Can Risk-Assessment Tools be Feasibly Used in the Health Benefit Marketplace?," *Advances in Health Economics and Health Services Research*, 12, JAI Press, 1991.	
JZ	13.	Goodman, Michael J. et al, "Persistence of Health Care Expense in an Insured Working Population," *Advances in Health Economics and Health Services Research*, 12, JAI Press, 1991.	
JZ	14.	Fuller, Douglas N., " AN EXPLORATION OF POPULATION CLASSIFICATION FOR MANAGED HEALTHCARE WITHIN A STATE-BASED MODELING FRAMEWORK" University of Virginia Dissertation, dated Jan. 2000, published May 22, 2000 (see attached letter from Director of Cataloging Services, University of Virginia Library).	

Examiner Signature	[signature]	Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional). [2] Applicant is to place a check mark here if English language Translation is attached.

Notice of References Cited	Application/Control No. 10/038,977	Applicant(s)/Patent Under Reexamination FULLER, DOUGLAS NEAL	
	Examiner Jerry Zhu	Art Unit 2121	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-			
	B	US-			
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Ferry Butar Butar, "Empirical Bayes Methods in Survey Sampling," August, 1997,
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

EMPIRICAL BAYES METHODS IN SURVEY SAMPLING

by

Ferry Butar Butar

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Mathematics and Statistics

Under the Supervision of Professor Parthasarathi Lahiri

Lincoln, Nebraska

August, 1997

DISSERTATION TITLE

Empirical Bayes Methods in Survey Sampling

BY

Ferry Butar Butar

SUPERVISORY COMMITTEE:

APPROVED	DATE
P. Lahiri Signature	6/20/97
Parthasarathi Lahiri Typed Name	
K.M. Lal Saxena Signature	6/20/97
K.M. Lal Saxena Typed Name	
Allan L. McCutcheon Signature	6-20-97
Allan L. McCutcheon Typed Name	
C. M. Ramsay Signature	6/24/97
Colin M. Ramsay Typed Name	
Signature	
Typed Name	
Signature	
Typed Name	

GRADUATE COLLEGE
UNIVERSITY OF NEBRASKA

EMPIRICAL BAYES METHODS IN SURVEY SAMPLING

Ferry Butar Butar, Ph.D.
University of Nebraska, 1997

Advisor: Parthasarati Lahiri

This dissertation concerns two problems in survey sampling: (a) small-area estimation and (b) estimation in finite population sampling. Both the topics have received considerable attention in recent years.

Empirical Bayes method has been found to be very useful in small area estimator and finite population sampling. The method is very effective in combining relevant information from the sample surveys, various administrative records and the census data.

The first half of the dissertation is devoted to small area estimation. In large scale national sample surveys, the sampling designs are determined so as to obtain reliable estimates of various characteristics of interest at the national level. Due to the availability of relatively small samples, the regular designed-based estimators perform poorly at the subnational level (e.g., state, county, etc.) when compared to the corresponding estimator at the national level. Similar situation arises when estimates are needed for a subgroup of the population obtained by classifying the population according to various demographic characteristics (e.g., age. race. sex, etc.). Such problems in survey sampling literature are known as small area estimation problems. Reliable small area statistics are needed in regional planning and in allocation of government resources.

The following research has been conducted in the small area estimation prob-

lems:

(a) A unified model is proposed which covers various specific small area models considered in the literature;

(b) A general measure of uncertainty of the proposed empirical Bayes estimator is considered and

(c) Small area estimation method under a random sampling variance model is developed.

Later part of the dissertation concerns empirical Bayes estimation of different stratum means and variances when samples are obtained using a stratified simple random sampling design. The method is effective specially when a moderately large samples are available from any given stratum. There are three main features of this research:

(a) In order to reduce the effect of overshrinking bias associated with the usual empirical Bayes procedures, stratum specific random effects are introduced through the sampling variances;

(b) General measures of uncertainty are proposed for the empirical Bayes point estimators of finite population means and variances;

(c) Laplace's second order approximation is used to approximate the one-dimensional integrals involved in the empirical Bayes point estimators and the measures of uncertainty of the point estimators. The approximation is specially helpful in obtaining the measures of uncertainty of the empirical Bayes estimators since the measures are based on Monte Carlo methods where checking the accuracy of the numerical integration method at each step of the replication is troublesome.

to the memory of my father

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Professor Parthasarathi Lahiri for being my major advisor and originally proposing the problems contained in the dissertation. I got his support and encouragement all through. Without his enormous patience, encouragement and guidence. it would not have been possible to complete this dissertation. I consider myself lucky to have him as my dissertation advisor.

I would like to thank Professors K.M. Lal Saxena, Allan L. McCutcheon and Colin M. Ramsay for serving on my committee.

I would also like to acknowledge Professor Arijit Chaudhuri of the Indian Statistical Institute for his assistance and support.

I share this achievement with my mother. my wife Rosmauly and my children Artha, Belinda, Cornelius for their constant love, patience and encouragement and dedicate it to the memory of my father.

CONTENTS

LIST OF TABLES

page

CHAPTER 1

INTRODUCTION

1.1 Literature Review

Empirical Bayes method as an area of research has received considerable importance in recent years due to its proven optimality properties, simplicity and wide range applications. Empirical Bayes estimators have been found to be very effective in small-area estimation and finite population sampling, the topics of this dissertation. According to this method, the Bayes estimator of the unknown parameter of interest is first obtained by using a suitable Bayesian model which combines information from various sources. The unknown parameters of the prior distribution are then estimated by a classical method such as method of moments. method of maximum likelihood (ML), method of residual maximum likelihood (REML), etc. The resulting estimator is the so-called empirical Bayes estimator.

Empirical Bayes method in the context of nonparametric estimation of a completely unspecified prior distribution was introduced by Robbins (1955). Efron and Morris (1973 , 1975) proposed an empirical Bayes method in the parametric setting. Morris (1983) gave an excellent account of the empirical approach and its applications.

An empirical Bayes estimator is usually obtained in a closed form and thus is appealing to the practitioners. One main criticism against the empirical Bayes method is that it does not provide a measure of uncertainty, specially for complex models, which captures all sources of variabilities. An estimator of the integrated Bayes risk of the Bayes estimator can be naively taken as a measure of uncertainty of an empirical Bayes estimator. But this measure can severely underestimate the true uncertainty of the empirical Bayes estimator since it does not incorporate the variabilities due to the estimation of various parameters of the prior distribution.

There have been several attempts to incorporate this extra uncertainty in the estimation procedure. Morris (1983) proposed a measure of uncertainty for his empirical Bayes estimator by approximating the posterior variance of a hierarchical Bayes solution which assumes flat priors on the hyperparameters. He, however, considered a simple one way balanced random effects model. Kass and Steffey (1989) considered Laplace method to approximate a hierarchical Bayes solution. Prasad and Rao (1990) obtained a measure of uncertainty of their EBLUP (same as empirical Bayes) by first approximating the mean squared error (MSE) by delta method and then estimating the approximated MSE. Their method is, however, restricted to the ANOVA method of variance component estimation. Recently, Datta and Lahiri (1997) unified the Prasad-Rao theory for a general mixed linear model and the theory is valid for many methods of variance component estimation including ML and REML. Lahiri and Rao (1995) extended the Prasad-Rao method for non-normal situation.

In this dissertation we shall develop empirical Bayes theory in small area estimation and finite population sampling.

1.1.1 Small Area Estimation

The sampling design and the sample size of most of the large scale national surveys are usually determined so as to produce reliable estimates of various characteristics of interest at the national level. Quite often there is a need to produce similar estimates at the subnational levels (e.g., states, counties, etc.). The direct survey method (see Cochran 1977) fails to provide reliable estimate for a subnational region due to small samples available from the region. Similar situation arises when estimates are needed for domains obtained by classifying the population according to various demographic characteristics (e.g., age, race, sex, etc.). Such problems in survey sampling literature are known as small area estimation problems. Reliable small area statistics are needed in regional planning and in allocation of government resources. The U.S. Census Bureau, the Bureau of the

Labor Statistics, the Statistics Canada, the Ministry of Planning, and other federal and local government agencies are interested in developing reliable small area statistics. See the working paper prepared by the subcommittee on small area estimation (1993) for a number of important small area estimation problems recently encountered by the U.S. federal agencies.

The history of small area statistics can be traced back to the 11th century in England (see Brackstone, 1987). Records of births. baptism, marriages. death. etc.. were used to produce various small area statistics. In those early days, sources of small area statistics were limited to various administrative records available from the local governments.

In the past fifty years sample survey has become an important component in many countries' statistical programs. Sample surveys have been very successful in supplying national and regional statistical data on a regular basis.

Due to budgetary constraints, it is not possible to collect adequate samples sizes from the small areas. When information on one or more relevant covariates is available, synthetic estimators, i.e., regression estimators, have been proposed in the small area literature (see Gonzales, 1973. Ericksen, 1974). Although the synthetic estimators have small variances compared to the direct survey estimators. they tend to be biased as they do not use the information on the characteristic of interest directly obtainable from the sample survey. A compromise between the direct survey and the synthetic estimation is the method of composite estimation (see Holt et al., 1979). Broadly defined, a composite estimator is a weighted average of a direct survey estimator and a synthetic estimator. The synthetic and composite estimators are usually obtained by *implicit* or *explicit* models which *borrow strength* from related resources.

Morrison (1971) described small area estimation methods which were used prior to 1970. Purcell and Kish (1979) reviewed demographic methods as well as statistical methods of estimation for small domains. National Research Council (1980) gave a detailed information as well as evaluation of the Census Bureau's procedure

for making post-censual estimates of the population and per-capita income for local areas. Zidek (1982) introduced a criterion that could be used to evaluate the relative performances of different methods for estimating the population of a small area and McCullagh and Zidek (1987) elaborated that criterion. Statistics Canada (1987) provided an overview and evaluation of the population estimation methods used in Canada. Schaible (1992) provided estimates on small area used in the U.S. Federal programs. For a review of the history of small area estimation. various small area estimation procedures and their applications, the reader is referred to Rao (1986), Chaudhuri (1992) and Ghosh and Rao (1994).

Due to the growing demand of small area estimation, many symposia and workshops on small area statistics were organized during the last two decades. The list of conferences and symposia includes National Institute of Drug Abuse, Princeton Conference (1979), International Symposium on Small Area Statistics, Ottawa (see Platek *et al.* (1987) for the invited papers and Platek and Singh (1986) for the contributed papers), International Symposium on Small Area Statistics, New Orleans. 1988, Workshop on Small-Area Estimation for Military Personnel Planning. Washington, D.C., 1989 and International Scientific Conference on Small-Area Statistics on Survey Design, Warsaw, Poland, 1992.

Empirical Bayes method has been extensively used in small area estimation and related problems. We now present a few specific application of empirical Bayes method in small area estimation and related problems.

Carter and Rolph (1974) considered estimation of the probabilities of false fire alarms reported from many street boxes of New York City. Using data collected from the street boxes during 1967-69, they provided empirical Bayes estimates of the probabilities of false alarms at various street boxes for the year 1970. In order to estimate the probability of a false alarm for a given box, they combined information from that box and all the boxes in the neighborhood.

Fay and Herriot (1979) generalized the Carter-Rolph model to incorporate information on a number of covariates and proposed empirical Bayes estimates of

per-capita incomes of small-places (population less than 1000). Their empirical Bayes estimator of per-capita income is a weighted average of the Current Population Survey estimator of the per-capita income and a regression estimator which utilizes tax return data for the year 1969 and the data on housing from the 1970 census.

Battese *et al.* (1988) considered the empirical best linear unbiased prediction (same as empirical Bayes) of areas under corn and soybeans for 12 counties in northern Iowa for the year 1978. They combined information from two sources. The direct information came from the 1978 June Enumeration Survey. The USDA Statistical Reporting Service field staff determined the areas under corn and soybeans in 37 sample segments (each segment is about 250 hectares) of 12 counties by interviewing farm operators. The second source of data was LANDSAT satellite data. Based on LANDSAT readings obtained during August and September of 1978, Battese *et al.* (1988) used the USDA procedures to classify the crop for all pixels (stands for *picture elements* extending over .45 hectares) in the 12 counties. A mixed linear regression model was then considered to establish a relationship between the survey and satellite data for the prediction purposes.

Every ten years the U.S. Census Bureau undertakes a census to account for its population. Unfortunately, the census counts have been found to be imperfect despite the fact that the Census Bureau puts a lot of efforts to make this massive project successful. There are various reasons for this imperfection of the census counts. Researchers found that the accuracy of the census counts depends quite a bit on various demographic factors (e.g., age, race, sex, etc.), owners and non-owners (renters) of dwellings and the geography. Since the census counts are used to apportion congressional seats and allocation of federal funds in various federal programs, the differential undercount poses a serious problem. We referred to a special issue of Survey Methodology (see Vol 18, 1992) which contains papers of Cressie and Datta *et al.* discussing various issues and methods to adjust for the census counts.

The department of Health and Human Services (HHS) uses estimates of the median income of four-person families at the state level to formulate its energy assistance program for low income families. Such data are provided by the U.S. Census Bureau for all the states and the District of Columbia on an annual basis. Currently the Census Bureau uses an empirical Bayes estimator based on the work by Fay (1987) (see also Fay *et al.* 1993; Datta *et al.* 1991; Datta *et al.* 1996; Ghosh *et al.* 1996). These papers used data from annual demographic supplement to the March sample of the Current Population Survey (CPS) which provides annual median income by states and family sizes, the decennial censuses and the Bureau of Economic Analysis which provides annual estimates of per-capita income for the states.

1.1.2 Finite Population Sampling

The results from the empirical Bayes estimation for small area characteristics can be adapted in estimating finite population means and variances from stratified simple random sampling. For a traditional design based approach to the finite population sampling, the reader is referred to Cohran (1977). Ericson (1969a) put forward an elegant formulation of the subjective Bayes approach to the finite population sampling. In his approach, he first assumed that the finite population is a realization from a hypothetical population which is the usual assumption in the super-population approach in finite population theory (see Royall 1970). At the second stage, Ericson (1969a) assumed a subjective prior distribution on the parameters of the super-population model. In practice, it is generally difficult to apply Ericson's Bayesian method since the prior parameters are hardly known. Ghosh and Meeden (1986) considered an empirical Bayes approach under a stratified simple random sampling, using an one-way random effects model. They successfully demonstrated that their method can be very effective in repeated surveys and small-area estimation. Their empirical Bayes estimator is asymptotically optimal in the sense of Robbins (1955). Later on Ghosh and Lahiri (1987) relaxed

the normality assumption of Ghosh and Meeden (1986) and showed that Ghosh-Meeden estimator is robust under the assumption of *posterior linearity* (see Ericson 1969 b; Goldstein 1975; Hartigan 1969). The Ghosh-Meeden empirical Bayes estimator can also be motivated from the best linear prediction approach of Prasad and Rao (1990). Nandram and Sedransk (1993) extended the Ghosh-Meeden estimator under different but random sampling variances. Recently, Arora *et al.* (1997) considered an alternative to the Nandram-Sedransk method. Their method can incorporate relevant auxiliary information which may be available from various administrative records and censuses. They also proposed, for the first time, a measure of uncertainty of the empirical Bayes estimator of finite population means which can incorporate uncertainty due to estimation of all the parameters in the Bayesian model. Their method is an extension of the parametric bootstrap method proposed earlier by Laird and Louis (1987) to the finite population sampling.

For the last fifteen years, there has been a growing demand from both the public and private sectors to produce reliable statistics for various subgroups of a finite population. According to Brakstone (1987) "there is in Canada, and probably in other countries too, an increasing government concern with issues of distribution . equity and disparity." Consider the problem of comparing the income distribution for various geographical areas of a country. Is it enough to consider just the per-capita income? Probably not. since two geographical areas may be comparable in terms of their per-capita incomes, yet they may vary considerably in terms of diversity which can be measured by the variances of their income distributions. Although the problem of finite population variances for different geographic groups is a very important problem. it has received relatively less attention than the problem of estimation of means, ratios and proportions for different geographical areas in the finite population sampling.

Ericson (1969a) briefly addressed the problem of the Bayesian estimation of a finite population variance under simple random sampling. Datta and Ghosh (1993) provided a unified approach to the Bayesian estimation of different strata

variances in finite population sampling under stratified random sampling. Ghosh and Lahiri (1987) considered the problem using a linear empirical Bayes approach. Lahiri and Tiwari (1990) proposed a nonparametric empirical Bayes estimation using the Dirichlet process prior (Ferguson 1973).

Note that the model considered by Datta and Ghosh (1993) does not incorporate stratum specific random effects through the scale components. Although, this synthetic assumption may have insignificant effect in the estimation of different stratum means, it may cause unduly shrinkage in the Bayes estimator of different stratum variances. Ghosh and Lahiri (1987) and Lahiri and Tiwari (1990) introduced random stratum effects through the scale parameters, but even then they failed to overcome the overshrinkage problem primary because of the linear nature of their Bayes estimators. However, we realize that the linear empirical Bayes procedure of Ghosh and Lahiri (1987) and the nonparametric empirical Bayes approach of Lahiri and Tiwari (1990) are very robust and it is difficult to resolve the problem associated with overshrinking without being specific about the distribution of the stratum specific random scale effects.

1.2 The Subject of This Dissertation

The organization of this dissertation is as follows:
In Chapter 2, we present a unified approach to empirical Bayes estimation in small area estimation and related problems. A simple measure of uncertainty which incorporates all sources of variations is proposed and hence the chapter addresses an outstanding problem in empirical Bayes estimation. Using a simple small area model, we show that our proposed measure of uncertainty enjoys both the frequentist and Bayesian properties. The method is also validated using real life examples and Monte Carlo simulations.

Small area estimation under random sampling variances has recently received a lot of attention. In Chapter 3, we consider the estimation of both small area

means and variances under a random sampling variances model. We use Laplace's second order approximation to obtain closed form formulae for the Bayes and empirical Bayes estimators. We demonstrate the accuracy of the approximations using numerical examples. In order to compare the performances of the proposed empirical Bayes estimators, we carry out Monte Carlo simulations.

We address empirical Bayes estimation in finite population sampling in Chapters 4, 5 and 6. In Chapters 4 and 5, we obtain useful approximations to the estimators earlier proposed by Arora (1994). In these chapters, we also consider the important problem of measuring the uncertainty of the proposed empirical Bayes estimators in finite population sampling. In Chapter 6, we consider empirical Bayes estimation of finite population proportions when samples are drawn using a stratified simple random sampling design.

CHAPTER 2

Estimation in a Mixed Linear Normal Model

2.1 Introduction

The main objective of this chapter is to develop a simple measure of uncertainty of an empirical Bayes small area estimator under a fairly general longitudinal mixed linear model. The proposed model covers many important small area models considered earlier in the literature including the Fay-Herriot and the nested error regression models described in section 2.2. In section 2.3, we introduce the general model and consider empirical Bayes point estimation of a general mixed effect. The empirical Bayes point estimator is identical with the empirical best linear unbiased predictor (EBLUP) proposed by Prasad and Rao (1990). Despite the popularity of the empirical Bayes method, the literature on empirical Bayes measure of uncertainty is not very rich as explained in Chapter 1. In section 2.4, we propose a unified measure of uncertainty for the proposed empirical Bayes estimator. Two real life examples are considered in section 2.5 to illustrate our method. Finally, in section 2.6, a Monte Carlo simulation is performed to validate the proposed measure of uncertainty of the empirical Bayes estimator. Proofs of all the theorems are given in the appendix to this chapter.

2.2 Two Models

2.2.1 Fay-Herriot Model

Fay and Herriot (1979) considered empirical Bayes method to estimate per-capita incomes of small places (population less than 1000). In order to obtain their empirical Bayes estimation, they used the following model: (i) $Y_i \mid \theta_i \overset{ind}{\sim} N(\theta_i, D_i)$, $i = 1, \ldots, m$; (ii) *Apriori*, $\theta_i \overset{ind}{\sim} N(x_i'\beta,\ A)$, $i = 1, \ldots, m$, where Y_i

is the survey estimator of per-capita income for the ith area, D_i is known the sampling variance of Y_i; $x_i = (x_{i1}, x_{i2}, ..., x_{ik})'$ is a vector of known benchmark variables obtainable from the 1969 tax return data and 1970 census data.

2.2.2 Nested-Error Regression Model

Battese *et al.* (1988) proposed a nested-error regression model to predict areas under corn and soybeans for 12 counties in Northern Iowa. They assumed that $y_{ij} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{2ij} + u_{ij}$, where i is a subscript for the county ($i = 1, ..., 12$) and j is a subscript for a segment within a given county (j=1,...,n_i, where n_i is the number of segments in the ith county). Here y_{ij} is the number of hectares under corn (soybeans) in the jth segment of the ith county as reported in the June Enumerative Survey; x_{1ij} is the number of pixels of corn and x_{2ij} is the number of pixels of soybeans for the jth segment in the ith county and β_0, β_1, β_2 are unknown parameters. The covariates x_{1ij} and x_{2ij} were obtained for the sample segments as well as for non sample segments using Landsat Satellite data. The random error u_{ij}, associated with the reported crop area y_{ij}, is expressed as $u_{ij} = v_i + e_{ij}$, where v_i is a random effect due to the ith county and e_{ij} is a pure error for the jth segment in the ith county. They assumed $v_i \overset{iid}{\sim} N(0, \sigma_v^2)$, e_{ij} ($j = 1, ..., n_i; i = 1, ..., 12$) $\overset{iid}{\sim} N(0, \sigma_e^2)$, $Cov(v_i, e_{ij}) = 0$, $Cov(v_i, v_{i'}) = 0$ if $(i \neq i')$ and $Cov(e_{ij}, e_{i'j'}) = 0$ if $(i, j) \neq (i', j')$.

Those two models discussed above are special cases of the general mixed linear model described in section 2.3.

2.3 Empirical Bayes Point Estimation

Let X_i and Z_i be n_i x p and n_i x k_i matrices of constants. Let $n = \sum_{i=1}^m n_i$ and $k = \sum_{i=1}^m k_i$. Consider the following Bayesian model:

Model 1:

(i) Conditional on a $k_i \times 1$ random vector U_i, Y_i's are independent with
$Y_i \mid U_i \sim N_{n_i}(X_i\beta + Z_i U_i,\ R_i), \quad i = 1, \ldots, m;$

(ii) *Apriori*, $U_i \stackrel{ind}{\sim} N_{k_i}(0, G_i), \quad i = 1, \ldots, m;$

where $R_i = R_i(\psi)$ and $G_i = G_i(\psi)$ are respectively $n_i \times n_i$ and $k_i \times k_i$ matrices which possibly depend on ψ, a $s \times 1$ vector of variance components.

Consider the estimation of $\theta_i = l_i' \beta + \lambda_i' U_i$, where l_i and λ_i are $p \times 1$ and $k_i \times 1$ vector of known constants respectively.

From Model 1, first, we need to find the distribution of $U_i \mid Y_i$. Now, in the joint density $f(Y, U)$ is

$$\propto exp -\frac{1}{2}\sum\left\{(Y_i - X_i\beta - Z_iU_i)'R_i^{-1}(Y_i - X_i\beta - Z_iU_i) + U_iG_i^{-1}U_i\right\} \quad (1)$$

To find the distribution of $U_i \mid Y_i$, look at the exponent of (1), terms involving U_i

$$\begin{aligned}
& \sum\left\{(Y_i - X_i\beta - Z_iU_i)'R_i^{-1}(Y_i - X_i\beta - Z_iU_i) + U_iG_i^{-1}U_i\right\} \\
= & \sum(Y_i - X_i\beta)'R_i^{-1}(Y_i - X_i\beta) - 2\sum(Y_i - X_i\beta)'R_i^{-1}Z_i U_i \\
+ & \sum U_i'Z_i'R_i^{-1}Z_iU_i + \sum U_iG_i^{-1}U_i \\
= & \sum(Y_i - X_i\beta)'R_i^{-1}(Y_i - X_i\beta) - 2\sum(Y_i - X_i\beta)'R_i^{-1}Z_i U_i \\
+ & \sum U_i'(G_i^{-1} + Z_i'R_i^{-1}Z_i)U_i \\
= & \sum(Y_i - X_i\beta)'R_i^{-1}(Y_i - X_i\beta) \\
+ & \sum\left\{U_i - (G_i^{-1} + Z_i'R_i^{-1}Z_i)^{-1}Z_i'R_i^{-1}(Y_i - X_i\beta)\right\}'(G_i^{-1} + Z_i'R_i^{-1}Z_i) \\
& \left\{U_i - (G_i^{-1} + Z_i'R_i^{-1}Z_i)^{-1}Z_i'R_i^{-1}(Y_i - X_i\beta)\right\} \\
- & \sum(Y_i - X_i\beta)'R_i^{-1}Z_i(G_i^{-1} + Z_i'R_i^{-1}Z_i)^{-1}Z_i'R_i^{-1}(Y_i - X_i\beta) \qquad (2)
\end{aligned}$$

Hence, the distribution of $U_i \mid Y_i$ is normal i.e.,

$$U_i \mid Y_i; \beta, \psi \sim N\left[(G_i^{-1} + Z_i'R_i^{-1}Z_i)^{-1}Z_i'R_i^{-1}(Y_i - X_i\beta), (G_i^{-1} + Z_i'R_i^{-1}Z_i)^{-1}\right]$$

Claim that $(G_i^{-1} + Z_i'R_i^{-1}Z_i)^{-1}Z_i'R_i^{-1} = G_iZ_i'V_i^{-1}$.

Proof:

$$\begin{aligned}
V_i^{-1} &= (R_i + Z_iG_iZ_i')^{-1} \\
&= R_i^{-1} - R_i^{-1}Z_i(Z_i'R_i^{-1}Z_i + G_i^{-1})^{-1}Z_i'R_i^{-1}
\end{aligned}$$

$$
\begin{aligned}
&= R_i^{-1} - R_i^{-1} Z_i \left[(Z_i' R_i^{-1} Z_i G_i + I_{k_i}) G_i^{-1} \right]^{-1} Z_i' R_i^{-1} \\
&= R_i^{-1} - R_i^{-1} Z_i G_i (Z_i' R_i^{-1} Z_i G_i + I_{k_i})^{-1} Z_i' R_i^{-1} \\
&= R_i^{-1} - R_i^{-1} Z_i G_i \Sigma Z_i' R_i^{-1}
\end{aligned}
$$

where $\Sigma = (Z_i' R_i^{-1} Z_i G_i + I_{k_i})^{-1}$, $\Sigma^{-1} = Z_i' R_i^{-1} Z_i G_i + I_{k_i}$, thus $Z_i' R_i^{-1} Z_i G_i = \Sigma^{-1} - I_{k_i}$, ($I_{k_i}$ being the identity matrix of order $k_i \times k_i$). Now,

$$
\begin{aligned}
Z_i' V_i^{-1} &= Z_i' R^{-1} - Z_i' R_i^{-1} Z_i G_i \Sigma Z_i' R_i^{-1} \\
&= Z_i' R_i^{-1} - (\Sigma^{-1} - I_{k_i}) \Sigma Z_i' R_i^{-1} \\
&= Z_i' R_i^{-1} - Z_i' R_i^{-1} \Sigma Z_i' R_i^{-1} \\
&= (Z_i' R_i^{-1} Z_i G_i + I_{k_i})^{-1} Z_i' R_i^{-1}
\end{aligned}
$$

then

$$
\begin{aligned}
G_i Z_i' V_i^{-1} &= G_i (Z_i' R_i^{-1} Z_i G_i + I_{k_i})^{-1} Z_i' R_i^{-1} \\
&= \left[(Z_i' R_i^{-1} Z_i G_i + I_{k_i}) G_i^{-1} \right]^{-1} Z_i' R_i^{-1} \\
&= \left[(Z_i' R_i^{-1} Z_i + G_i^{-1}) \right]^{-1} Z_i' R_i^{-1}
\end{aligned}
$$

Then, the probability distribution of $U_i \mid Y_i$ can be written as

$$
U_i \mid Y_i \sim N \left[G_i\, Z_i' V_i^{-1} (Y_i - X_i \beta), (G_i^{-1} + Z_i' R_i^{-1} Z_i)^{-1} \right]
$$

Under the above model and squared error loss function, the Bayes estimator of θ_i is given by

$$
\begin{aligned}
\hat{\theta}_i^{B} &= E[\theta_i \mid Y_i; \beta, \psi] = l_i' \beta + \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) (Y_i - X_i \beta) \\
&= \hat{\theta}_i(Y_i; \beta, \psi), say. \qquad (3)
\end{aligned}
$$

where $V_i(\psi) = R_i + Z_i G_i Z_i'$ $(i = 1, \ldots, m)$.
When ψ is known, but β is unknown, β is estimated by the maximum likelihood estimator $\hat{\beta}(\psi)$, where $\hat{\beta}(\psi) = \left[\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i \right]^{-1} \left[\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) Y_i \right]$. Plugging in $\hat{\beta}(\psi)$ for β in $\hat{\theta}_i^{B}$, we get the following empirical Bayes estimator of θ_i:

$$
\tilde{\theta}_i^{EB} = \hat{\theta}_i(Y_i; \hat{\beta}(\psi), \psi) = l_i' \hat{\beta}(\psi) + \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) [Y_i - X_i \hat{\beta}(\psi)]. \quad (4)
$$

Note that $\tilde{\theta}_i^{EB}$ is a robust estimator in the sense that it can be viewed as a best linear unbiased predictor (see Prasad and Rao 1990) or a linear empirical Bayes estimator (Ghosh and Lahiri 1987).

In practice β and ψ are both unknown. In this case, an empirical Bayes estimator of θ_i is obtained as $\hat{\theta}_i^{EB} = \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi})$, where $\hat{\psi}$ is an estimator of ψ which satisfies the regularity conditions of Datta and Lahiri (1997) given in the Appendix. We shall assume that $E(\hat{\psi} - \psi) = -m^{-1}B(\psi) + o(m^{-1})$, where the functional form of $B(\psi)$ is known.

Example: Model 1 covers the Fay-Herriot example in section 2.2.1.

Fay-Herriot Model (see Fay and Herriot 1979)

(i) $Y_i \mid \theta_i \overset{ind}{\sim} N(x_i'\beta + \theta_i, D_i), \quad i = 1, \ldots, m;$

(ii) *Apriori*, $\theta_i \overset{ind}{\sim} N(0, A), i = 1, \ldots, m;$

where D_i's are known and x_i's are p x 1 vector of known constants. In the notations of the Model 1, $n_i = k_i = 1$, $Z_i = 1$, $U_i = \theta_i$, $\psi = A, R_i(\psi) = D_i$ and $G_i(\psi) = A \ \ (i = 1, \cdots, m)$.

In this case, we may consider the following empirical Bayes estimator of θ_i:

$$\hat{\theta}_i^{EB} = x_i'\hat{\beta} + (1 - \hat{B}_i)(Y_i - x_i'\hat{\beta}),$$

where $\hat{\beta} = \left(\sum_{i=1}^m \frac{1}{\hat{A} + D_i} x_i x_i'\right)^{-1} \left(\sum_{i=1}^m \frac{1}{\hat{A} + D_i} x_i Y_i\right)$,

$$\hat{A} = max\left(0, \tilde{A} = \frac{1}{m-p}\left[\sum_{i=1}^m (Y_i - x_i'\hat{\beta}_{ols})^2 - \sum_{i=1}^m (1 - h_{ii}) D_i\right]\right),$$

$\hat{\beta}_{ols} = (\sum_{i=1}^m x_i x_i')^{-1} (\sum_{i=1}^m x_i Y_i)$, $h_{ii} = x_i'(\sum_{i=1}^m x_i x_i')^{-1} x_i$, $\hat{B}_i = \frac{D_i}{\hat{A} + D_i}$. In this case, $B(\psi) = 0$.

For other particular cases of Model 1, see Carter and Rolph (1974), Fay and Herriot (1979), Battese *et al.* (1988), Lahiri and Wang (1992), among others.

2.4. A Measure of Uncertainty of $\hat{\theta}_i^{EB}$

Define the integrated Bayes risk of $\hat{\theta}_i^{EB}$ as $r(\hat{\theta}_i^{EB}) = E(\hat{\theta}_i^{EB} - \theta_i)^2$, where the expectation is taken with respect to Model 1. Note that $r(\hat{\theta}_i^{EB})$ is identical to the

MSE of $\hat{\theta}_i^{EB}$ as defined in Prasad and Rao (1990). A measure of uncertainty of $\tilde{\theta}_i^{B}$ is given by

$$
\begin{aligned}
r(\tilde{\theta}_i^{EB}) &= E(\tilde{\theta}_i^{EB} - \theta_i)^2 \\
&= E\left\{l_i'\hat{\beta}(\psi) + \lambda_i'\hat{U}_i - l_i'\beta - \lambda_i'U_i\right\}^2 \\
&= E\left\{l_i'(\hat{\beta}(\psi) - \beta) + \lambda_i'(\hat{U}_i - U_i)\right\}^2 \\
&= Var\left[(l_i'(\hat{\beta}(\psi) - \beta)\right] + Var\left[\lambda_i'(\hat{U}_i - U_i)\right] \\
&\quad + 2\, l_i'\, Cov\left\{(l_i'(\hat{\beta}(\psi) - \beta), \lambda_i'(\hat{U}_i - U_i)\right\} \\
&= l_i'[\sum_{i=1}^{m} X_i'\, V_i^{-1}(\psi)\, X_i]^{-1} l_i + \lambda_i' Var(\hat{U}_i)\lambda_i \\
&\quad + \lambda_i' Var(U_i)\lambda_i - 2\, \lambda_i'\, Cov(\hat{U}_i, U_i')\lambda_i \\
&\quad + 2\, l_i' Cov(\hat{\beta}, \hat{U}_i')\lambda_i - 2\lambda_i'\, Cov(\hat{\beta}, U_i')\lambda_i. \qquad (5)
\end{aligned}
$$

Note that $\hat{U}_i = G_i(\psi)\, Z_i'\, V_i^{-1}(\psi)$. The third term of (5) is $\lambda_i' G_i(\psi)\lambda_i$. Now, the second term of (5) is

$$
\begin{aligned}
Var(\hat{U}_i) &= Var\left\{G_i(\psi)Z_i'V_i^{-1}(\psi)(Y_i - X_i\hat{\beta}(\psi))\right\} \\
&= Var\left\{G_i(\psi)Z_i'V_i^{-1}(\psi)\{Y_i - X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)Y_i]\}\right\} \\
&= Var\{G_i(\psi)Z_i'V_i^{-1}(\psi)Y_i\} \\
&\quad + Var\{G_i(\psi)Z_i'V_i^{-1}(\psi)X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)Y_i\} \\
&\quad - 2\, Cov\left\{G_i(\psi)Z_i'V_i^{-1}(\psi)Y_i\}X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)Y_i].\right. \\
&\quad \left. Y_i'V_i(\psi)Z_iG_i(\psi)\right\} \\
&= G_i(\psi)Z_i'V_i^{-1}(\psi)Y_iZ_iG_i(\psi) \\
&\quad + Var\left\{G_i(\psi)Z_i'V_i^{-1}(\psi)X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}(X_1'V_1^{-1}(\psi)Y_1 + \ldots\right. \\
&\quad \left. + X_m'V_m^{-1}(\psi)Y_m)\right\} \\
&\quad - 2\, Cov\left\{G_i(\psi)Z_i'V_i^{-1}(\psi)Y_i\}X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}\right. \\
&\quad \left. X_i'V_i^{-1}(\psi)Y_i, Y_i'V_i(\psi)Z_iG_i(\psi)\right\} \\
&= G_i(\psi)Z_i'V_i^{-1}(\psi)Y_iZ_iG_i(\psi)
\end{aligned}
$$

$$+ \; G_i(\psi)Z_i'V_i^{-1}(\psi)X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}X_i'V_i^{-1}(\psi)Z_iG_i(\psi)$$

$$-2 \; G_i(\psi)Z_i'V_i^{-1}(\psi)X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}X_i'V_i^{-1}(\psi)Z_iG_i(\psi)$$

$$= \; G_i(\psi)Z_i'V_i^{-1}(\psi)\{I_{n_i} \; - \; X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}X_i'V_i^{-1}(\psi)\}Z_iG_i(\psi) \quad (6)$$

The fourth term of (5) becomes

$$\begin{aligned}
Cov\,(\hat{U}_i, U_i') &= Cov\left\{G_i(\psi)Z_i'V_i^{-1}(\psi)(Y_i - X_i\hat{\beta}(\psi)), U_i'\right\} \\
&= Cov\,\Big\{G_i(\psi)Z_i'V_i^{-1}(\psi)\,[X_i\beta + Z_iU_i + e_i \\
&\quad - X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}\{\sum_{i=1}^{m} X_i'\; V_i^{-1}(\psi)(X_i\beta + Z_iU_i + e_i)\}\Big]\,, U_i'\Big\} \\
&= Cov\,\Big\{G_i(\psi)Z_i'V_i^{-1}(\psi)\,[X_i\beta \; + \; Z_iU_i \; + \; e_i \\
&\quad - X_i\beta \; - \; X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)Z_iU_i] \\
&\quad - X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)e_i]\Big]\,, U_i'\Big\} \\
&= Cov\,\Big\{G_i(\psi)Z_i'V_i^{-1}(\psi)[Z_iU_i \; - \; X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1} \\
&\quad X_i'V_i^{-1}(\psi)Z_iU_i,\; U_i'\Big\} \\
&= Cov\,\Big\{G_i(\psi)Z_i'V_i^{-1}(\psi)[I_{n_i} \; - \; X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1} \\
&\quad X_i'V_i^{-1}(\psi)]Z_iU_i,\; U_i'\Big\} \\
&= G_i(\psi)Z_i'V_i^{-1}(\psi)[I_{n_i} \; - \; X_i[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1} \\
&\quad X_i'V_i^{-1}(\psi)]Z_iG_i(\psi) \qquad (7)
\end{aligned}$$

Since in the above $e_i's$ and $U_i's$ are independent and $U_i's$ are also independent. The fifth term of (5) is

$$\begin{aligned}
Cov\,(\hat{\beta},\; \hat{U}_i') &= Cov\,\left\{\hat{\beta}, [G_i(\psi)Z_i'V_i^{-1}(\psi)(Y_i - X_i\hat{\beta}(\psi))]'\right\} \\
&= Cov\,\{\hat{\beta}, Y_i'V_i^{-1}(\psi)Z_iG_i(\psi)\} \\
&\quad - Cov\,\{\hat{\beta}, \hat{\beta}'(\psi)X_iV_i^{-1}(\psi)Z_iG_i(\psi))\} \\
&= [\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i]^{-1}X_i'V_i^{-1}(\psi)Z_iG_i(\psi) \quad \textit{similar to the third term}
\end{aligned}$$

$$
\begin{aligned}
& - [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \\
= \ & 0 \qquad (8)
\end{aligned}
$$

The sixth term of (5) is

$$
\begin{aligned}
Cov\ (\hat{\beta},\ U_i') &= Cov \left\{ [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) Y_i], U_i' \right\} \\
&= Cov \left\{ [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} \sum_{i=1}^{m} X_i' V_i^{-1}(\psi)[X_i\beta + Z_i U_i + e_i], U_i' \right\} \\
&= Cov \left\{ [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i U_i, U_i' \right\} \\
&= [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \qquad (9)
\end{aligned}
$$

Now, using (6) $-$ (9), (5) becomes

$$
\begin{aligned}
r(\tilde{\theta}_i^{EB}) &= l_i' [\sum_{i=1}^{m} X_i'\ V_i^{-1}(\psi)\ X_i]^{-1} l_i \ + \ \lambda_i' G_i(\psi) \lambda_i \\
& \quad + \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) \{ I_{n_i} - X_i [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) \} Z_i G_i(\psi) \lambda_i \\
& \quad -2\ \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) \{ I_{n_i} - X_i [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) \} Z_i G_i(\psi) \lambda_i \\
& \quad -2\ l_i' [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i \\
&= \lambda_i' G_i(\psi) \lambda_i \ - \ \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) \\
& \quad \{ I_{n_i} \ - \ X_i [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) \} Z_i G_i(\psi) \lambda_i \\
& \quad + \ l_i' [\sum_{i=1}^{m} X_i'\ V_i^{-1}(\psi)\ X_i]^{-1} l_i \\
& \quad -2\ l_i' [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i \\
&= \lambda_i' G_i(\psi) \lambda_i \ - \ \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i \\
& \quad + \lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi) X_i [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i \\
& \quad + \ l_i' [\sum_{i=1}^{m} X_i'\ V_i^{-1}(\psi)\ X_i]^{-1} l_i \\
& \quad -2\ l_i' [\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i]^{-1} X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i
\end{aligned}
$$

$$
\begin{aligned}
&= \lambda_i' \, G_i(\psi) \, [I_{k_i} \; - \; Z_i' \, V_i^{-1}(\psi) \; Z_i \; G_i(\psi)] \; \lambda_i \\
&\quad + (l_i - X_i' V_i^{-1}(\psi) Z_i) G_i(\psi) \lambda_i)' (\sum_{i=1}^{m} X_i V_i^{-1}(\psi) X_i)^{-1} \\
&\quad (l_i - X_i' V_i^{-1}(\psi) Z_i) G_i(\psi) \lambda_i) \\
&= g_{1i}(\psi) \; + \; g_{2i}(\psi), \; say. \qquad (10)
\end{aligned}
$$

where $g_{1i}(\psi) = \lambda_i' \, G_i(\psi) \, [I_{k_i} - Z_i' \, V_i^{-1}(\psi) \; Z_i \; G_i(\psi)] \; \lambda_i$, ($I_{k_i}$ being the identity matrix of order $k_i \times k_i$), $g_{2i}(\psi) = (l_i - X_i' \, V_i^{-1}(\psi) \; Z_i \; G_i \; (\psi)\lambda_i)'(\sum_{i=1}^{m} X_i' \, V_i^{-1}(\psi) \; X_i)^{-1}$ $(l_i \; - \; X_i' \; V_i^{-1}(\psi) \; Z_i \; G_i(\psi) \; \lambda_i)$, (see Kackar and Harville 1984; Prasad and Rao 1990). An appropriate estimator of $r(\hat{\theta}_i^{EB})$ can be taken as a measure of uncertainty of $\hat{\theta}_i^{EB}$. First note that

$$
r(\hat{\theta}_i^{EB}) \; = \; g_{1i}(\psi) + g_{2i}(\psi) + \; E[\hat{\theta}_i^{EB} \; - \; \tilde{\theta}_i^{EB}]^2, \qquad (11)
$$

Note that $g_{1i}(\psi)$ is the measure of uncertainty of the Bayes estimator $\hat{\theta}_i^{B}$, $g_{2i}(\psi)$ is the uncertainty due to estimation of β and the third term is due to the estimation of ψ. One may naively approximate $r(\hat{\theta}_i^{EB})$ by $g_{1i}(\psi) + g_{2i}(\psi)$ which ignores the uncertainty due to estimation of ψ. Datta and Lahiri (1997) showed that under regularity conditions (RC),

$$
E(\hat{\theta}_i^{EB} - \tilde{\theta}_i^{EB})^2 = g_{3i}(\psi) \; + \; o(m^{-1}), \qquad (12)
$$

where $g_{3i}(\psi) = trace \; [L_i(\psi) V_i(\psi) L_i'(\psi) \sum(\psi)]$, $L_i(\psi) = \; col_{1 \le j \le s} \; L_{ij}'(\psi)$, $L_{ij}'(\psi) = \frac{\partial}{\partial \psi_j} \; (\lambda_i' G_i(\psi) Z_i' V_i^{-1}(\psi))$ and $\sum(\psi) = E(\hat{\psi} \; - \; \psi)(\hat{\psi} \; - \; \psi)'$. The expression for $\sum(\psi)$ for some standard methods of estimation of ψ (e.g., ANOVA, ML, REML) are given in Prasad and Rao (1990) and Datta and Lahiri (1997). Thus, the naive approximation, *i.e.*, $g_{1i}(\psi) \; + \; g_{2i}(\psi)$, could lead to a serious underestimation since $g_{3i}(\psi)$ is of order $O(m^{-1})$, same as the order of $g_{2i}(\psi)$. For this reason, in this paper we shall not ignore any term of order $O(m^{-1})$. A naive estimator of $r(\hat{\theta}_i^{EB})$ is obtained from the naive approximation and is given by $V_i^N \; = \; g_{1i}(\hat{\psi}) + g_{2i}(\hat{\psi})$. It follows from Datta and Lahiri (1997) that under regularity conditions (RC),

$$
E\left\{g_{1i}(\hat{\psi}) + g_{2i}(\hat{\psi})\right\} = g_{1i}(\psi) + g_{2i}(\psi) - m^{-1} B'(\psi) \nabla g_{1i}(\psi) - g_{3i}(\psi) + o(m^{-1}). \qquad (13)
$$

where $\nabla g_{1i}(\psi) = col_{1\le j\le s} \frac{\partial}{\partial\psi_j} g_{1i}(\psi)$, and estimated $g_{1i}(\psi)+g_{2i}(\psi)$ in (11) by $g_{1i}(\hat\psi)+g_{2i}(\hat\psi) + m^{-1}B'(\hat\psi)\nabla g_{1i}(\hat\psi) + g_{3i}(\hat\psi)$. where $B(\hat\psi)$, $\nabla g_{1i}(\hat\psi)$ and $g_{3i}(\hat\psi)$. are obtained from $B(\psi)$, $\nabla g_{1i}(\psi)$,and $g_{3i}(\psi)$ respectively when ψ is replaced by $\hat\psi$.

To estimate $r(\hat\theta_i^{EB})$ in (11), we shall estimate $g_{1i}(\psi)+g_{2i}(\psi)$ by the Datta-Lahiri or Prasad-Rao method. However, our approach differs from the Prasad-Rao or Datta-Lahiri method in the estimation of the third term of (11). To introduce the method consider the following bootstrap model:

Model 2:

(i) $Y_i^* \mid U_i^* \overset{ind}{\sim} N_{n_i}(X_i\hat\beta(\hat\psi)+Z_iU_i^*,\ \hat R_i), \quad i=1,\dots,m;$

(ii) *Apriori*, $U_i^* \overset{ind}{\sim} N_{k_i}(0,\ \hat G_i), \quad i=1,\dots,m;$

where $\hat R_i = R_i(\hat\psi)$ and $\hat G_i = G_i(\hat\psi)$.

A reasonable estimator of the third term of (11) is $E_*[\hat\theta_i(Y_i;\hat\beta^*(\hat\psi^*)\ ,\hat\psi^*) - \hat\theta_i(Y_i;\hat\beta^*(\hat\psi)\ ,\hat\psi)\ ]^2$, where E_* is the expectation with respect to Model 2. $\hat\beta^*(\hat\psi^*) = \left(\sum_{i=1}^m X_i' V_i^{-1}(\hat\psi^*)\ X_i\right)^{-1}\left(\sum_{i=1}^m X_i' V_i^{-1}(\hat\psi^*)\ Y_i^*\right)$ and the calculation of $\hat\psi^*$ is the same as that of $\hat\psi$ except that it is based on Y_i^*'s instead of Y_i's. It is shown in Theorem A.1 that:

$$E_*[\hat\theta_i(Y_i;\hat\beta^*(\hat\psi^*)\ ,\hat\psi^*) - \hat\theta_i(Y_i;\hat\beta^*(\hat\psi)\ ,\hat\psi)\ ]^2 = g_{4i}(\hat\psi;Y_i) + o_p(m^{-1}) \qquad (14)$$

where $g_{4i}(\hat\psi;Y_i) = trace[L_i(\hat\psi)\ [Y_i - X_i\hat\beta(\hat\psi)][Y_i - X_i\hat\beta(\hat\psi)]'L_i'(\hat\psi)\ \sum(\hat\psi)]$.

Thus, we now propose the following estimator of $r(\hat\theta_i^{EB})$:

$$V_i^P = g_{1i}(\hat\psi) + m^{-1}B'(\hat\psi)\nabla g_{1i}(\hat\psi) + g_{2i}(\hat\psi) + g_{3i}(\hat\psi) + g_{4i}(\hat\psi;Y_i). \qquad (15)$$

Laird and Louis (1987) proposed a measure of uncertainty of an empirical Bayes estimator for a very special case of Model 1, specifically the Fay-Herriot model with $x_i'\beta=\mu$ and $D_i = D$ $(i=1,\cdots,m)$. For the general model , one may extend their measure as $V_i^{LL} = E_*\left[g_{1i}(\hat\psi^*)\right] + V_*\left[\hat\theta_i(Y_i;\ \hat\beta^*(\hat\psi^*),\hat\psi^*)\right]$, where E_* and V_* are the expectation and the variance with respect to Model 2.

In Theorem A.2, we showed that under Model 1 and regularity conditions (RC).

$$E[V_i^{LL}] = r(\hat\theta_i^{EB}) - 2m^{-1}B'(\psi)\nabla g_{1i}(\psi) - g_{3i}(\psi) + o(m^{-1}). \qquad (16)$$

but

$$E[V_i^P] = r(\hat{\theta}_i^{EB}) + o(m^{-1}). \qquad (17)$$

Thus, unlike V_i^P, V_i^{LL} could lead to an underestimation of $r(\hat{\theta}_i^{EB})$ since the order of bias for V_i^{LL} is $O(m^{-1})$.

Remark 1. Consider a special case of the Fay-Herriot model when $x_i'\beta = \mu$ and $D_i = D$ $(i = 1, \cdots, m)$. An empirical Bayes estimator of θ_i is given by $\hat{\theta}_i^{EB} = \overline{Y} + (1 - \hat{B}_1)(Y_i - \overline{Y})$, where $\overline{Y} = \frac{1}{m}\sum Y_i$, $\hat{B}_1 = min(1.0, \hat{B})$ and $\hat{B} = \frac{D\,(m-1)}{\sum (Y_i - \overline{Y})^2}$. (see Prasad and Rao 1990; Laird and Louis 1987). Using a flat improper prior distributions on μ and $B = \frac{D}{A+D}$, Morris (1983a, 83b) suggested an approximation of the posterior variance of θ_i as a measure of uncertainty of $\hat{\theta}_i^{EB}$. See Kass and Steffey (1989) for alternative Bayesian approach, using the Laplace approximation. The measure is given by

$$V_i^M = (1 - \hat{B}_1)\,D + \frac{D\,\hat{B}_1}{m} + \frac{2\,\hat{B}_1^2}{m-3}(Y_i - \overline{Y})^2. \qquad (18)$$

To get the proposed measure of uncertainty, we shall solve $r(\hat{\theta}_i^{EB})$ i.e.,

$$r(\hat{\theta}_i^{EB}) = E_*[\hat{\theta}_i^{EB} - \theta_i]^2 = E[\tilde{\theta}_i^{EB} - \theta_i]^2 + E_*[\hat{\theta}_i^{EB} - \tilde{\theta}_i^{EB}]^2 \qquad (19)$$

The first term of (19) is

$$\begin{aligned}
E[\tilde{\theta}_i^{EB} - \theta_i]^2 &= E\,[\overline{Y} + (1 - B_1)\,(Y_i - \overline{Y}) - \theta_i]^2 \\
&= E\left\{(\overline{Y} - \mu) + (1 - B_1)[(Y_i - \mu) - (\overline{Y} - \mu)] - (\theta_i - \mu)\right\}^2 \\
&= E\left\{B_1(\overline{Y} - \mu) + (1 - B_1)(Y_i - \mu) - (\theta_i - \mu)\right\}^2 \\
&= B_1^2 E(\overline{Y} - \mu)^2 + (1 - B_1)^2 E(Y_i - \mu)^2 + (\theta_i - \mu)^2 \\
&\quad + 2B_1(1 - B_1)E(\overline{Y} - \mu)(Y_i - \mu) \\
&\quad - 2\,B_1 E(\overline{Y} - \mu)(\theta_i - \mu) - 2(1 - B_1)E(Y_i - \mu)(\theta_i - \mu) \\
&= B_1^2\frac{D}{mB_1} + (1 - B_1)^2\frac{D}{m} + A + 2B_1(1 - B_1)\frac{D}{mB_1} \\
&\quad - 2B_1 E\{(\theta_i - \mu)E(\overline{Y} - \mu) \mid \theta_i\} \\
&\quad - 2(1 - B_1)E\{(\theta_i - \mu)E(Y_i - \mu) \mid \theta_i\} \\
&= \frac{B_1 D}{m} + \frac{A^2}{A + D} + A + \frac{2}{m}\frac{AD}{A + D} - 2B_1\frac{A}{m} - 2(1 - B_1)A
\end{aligned}$$

$$
\begin{aligned}
&= \frac{B_1 D}{m} + \frac{A^2}{A+D} + A + \frac{2AD}{m(A+D)} - \frac{2AD}{m(A+D)} - \frac{2A^2}{A+D} \\
&= \frac{B_1 D}{m} + A - \frac{A^2}{A+D} \\
&= D(1-B_1) + \frac{B_1 D}{m} \qquad (20)
\end{aligned}
$$

Note that the last term of (20) are estimated by $D(1-\hat{B}_1) + \frac{2\hat{B}_1 D}{m} + \frac{m-3}{m-1}\frac{\hat{B}_1 D}{m}$. The second term of (19) is

$$
\begin{aligned}
E_*[\hat{\theta}_i^{EB} - \tilde{\theta}_i^{EB}]^2 &= E_*[\overline{Y} - (1-\hat{B}_1^*)(Y_i - \overline{Y}) - \overline{Y} - (1-\hat{B}_1)(Y_i - \overline{Y})]^2 \\
&= E_*[-(\hat{B}_1^* - \hat{B}_1)(Y_i - \overline{Y})]^2 \\
&= (Y_i - \overline{Y})^2\, E_*(\hat{B}_1^* - \hat{B}_1)^2 \\
&= (Y_i - \overline{Y})^2\, E_*(\hat{B}_1^*)^2 + \hat{B}_1^2 - \hat{B}_1^*\hat{B}_1 \\
&= (Y_i - \overline{Y})^2\, E_*\{(m-1)\hat{B}_1\frac{1}{u}\}^2 + \hat{B}_1^2 - 2\hat{B}_1\, E_*\{(m-1)\hat{B}_1\frac{1}{u}\} \\
&= (Y_i - \overline{Y})^2\, \frac{2\hat{B}_1}{m-5}, \qquad (21)
\end{aligned}
$$

where u in the above is chi-square with $m-1$ degree of freedom. Using (20) – (21) then $r(\hat{\theta}_i^{EB})$ or the proposed measure of uncertainty

$$
V_i^P = (1-\hat{B}_1)D + \frac{2D\hat{B}_1}{m} + \frac{m-3}{m-1}\frac{B_1}{m} + \frac{2\hat{B}_1^2}{m-5}(Y_i - \overline{Y})^2, \qquad (22)
$$

Similarly the Prasad-Rao (1990) measure of uncertainty is given by

$$
V_i^{PR} = (1-\hat{B}_1)\, D + \frac{D\,\hat{B}_1}{m} + \frac{4\,D\,\hat{B}_1}{m}, \qquad (23)
$$

The measure of uncertainty of $\hat{\theta}_i^{EB}$ using the method of Kass and Steffey (1989) is given by:

$$
V_i^{KSI} = (1-\hat{B}_1)\, D + \frac{D\,\hat{B}_1}{m} + \frac{2\,\hat{B}_1^2}{m-1}(Y_i - \overline{Y})^2, \qquad (24)
$$

One can also adjust the Kass-Steffey measure of uncertainty for order $O(m^{-1})$ bias and get

$$
V_i^{KSII} = V_i^{KSI} + \frac{2D\hat{B}_1}{m} \qquad (25)
$$

Laird and Louis (1987) proposed the following measure of uncertainty of $\hat{\theta}_i^{EB}$:

$$V_i^{LL} = (1 - \hat{B}_1)\, D + \frac{m-1}{m-5}\,\frac{D\,\hat{B}_1}{m} + \frac{2\,\hat{B}_1^2}{m\,-\,5}\,(Y_i\,-\,\overline{Y})^2. \qquad (26)$$

Thus, V_i^P and V_i^M are identical upto order $o_p(m^{-1})$. However, the difference between V_i^M and V_i^{LL} is of order $O_p(m^{-1})$. The Prasad-Rao measure V_i^{PR} cannot match a hierarchical Bayes solution since a hierarchical Bayes solution must be of the form $(1\,-\,E(B \mid Y))\,D\,+\,\frac{D\,E(B|Y)}{m}\,+\,(Y_i\,-\,\overline{Y})^2\,V(B \mid Y)$, where $E(B \mid Y)$ and $V(B \mid Y)$ are the posterior mean and variance of B, under a suitable prior on the hyper-parameters μ and B. We emphasize that the Prasad-Rao method gives exactly the same measure of uncertainty for all the small-areas unlike the other methods since the Prasad-Rao method does not depend on the individual Y_i.

Remark 2. It is shown in Theorem A.1 that $V_i^{LL}\,=\,g_{1i}(\hat{\psi})\,+\,g_{2i}(\hat{\psi})\,+\,g_{4i}(\hat{\psi};Y_i)$ $-\,m^{-1}B'(\hat{\psi})\nabla g_{1i}(\hat{\psi}) -\, g_{3i}(\hat{\psi})\,+\,o_p(m^{-1})$. Since $E\,g_{4i}(\hat{\psi};Y_i)\,=\,g_{3i}(\psi)\,+\,o(m^{-1})$, it is quite possible that for some i, V_i^{LL} could give us a measure which is less than the naive measure V_i^N (at least for large m).

For general case of the Fay-Herriot model, the proposed measure of uncertainty of $\hat{\theta}_i^{EB}$ is given by

$$\begin{aligned} V_i^P \;&=\; (1-\hat{B}_i)D_i + \hat{B}_i^2\, x_i'(\textstyle\sum D_i^{-1}\hat{B}_i x_i x_i')^{-1}x_i \\ &\quad + \frac{\hat{B}_i^3}{D_i}\frac{2}{m^2}\Sigma \hat{B}_i^{-2}D_i^2 + 2\frac{\hat{B}_i^4}{D_i^2}\frac{1}{m^2}\left(\Sigma D_i^2\hat{B}_i^{-2}\right)(Y_i - x_i'\hat{\beta})^2. \end{aligned} \qquad (27)$$

Note that in this case, the Laird-Louis method can be approximated by

$$\begin{aligned} V_i^P \;&=\; (1-\hat{B}_i)D_i + \hat{B}_i^2\; x_i'(\textstyle\sum D_i^{-1}\hat{B}_i x_i x_i')^{-1}x_i \\ &\quad + 2\frac{\hat{B}_i^4}{D_i^2}\frac{1}{m^2}\left(\Sigma D_i^2\hat{B}_i^{-2}\right)(Y_i - x_i'\hat{\beta})^2. \end{aligned} \qquad (28)$$

2.5. Two Examples

Efron and Morris (1975) successfully demonstrated the superiority of the empirical Bayes estimator (see section 3) over the classical estimator Y_i using the

famous baseball data which contains the batting averages of 18 major league baseball players. It is instructive to compare various measures of uncertainty of their empirical Bayes estimator. Table 1 presents various measures of uncertainty given in Remark 1. Amount of inflation of the measures which incorporate the uncertainty due to estimation of A is substantial when compared with the naive measure V_i^N. Note that V_i^{PR} and V_i^N are constant (0.341) and (0.153), respectively for all the baseball players. All the other measures of uncertainty change from player to player since they depend on the individual Y_i.

For the above example, $m = 18$ may be considered to be small. We now consider another example where $m = 51$ is moderately large. The U.S. Department of Health and Human Services (HHS) uses estimates of median income of four-person families at the state level to formulate its energy assistance program to low income families. Such data are provided by the U.S. Census Bureau for all the states (including the District of Columbia) on an annual basis. The current estimates are produced by an empirical Bayes procedure (see Fay 1987; Fay *et al.* 1993; Datta *et al.* 1996; Ghosh *et al.* 1996). The data we analyze provide the usual design-based estimates of 4-person families (Y_i) and its sampling variances D_i for all the 50 states and the District of Columbia for the year 1988. As a demonstration, in order to produce empirical Bayes estimates, using the Fay-Herriot model, we choose 1979 census estimates of median income of four-person families updated by the change in per-capita income obtainable from the Bureau of Economic Analysis. Our focus here is on the comparison of different measures of uncertainty of the empirical Bayes estimator. Table 2 presents the standard error of Y_i (i.e., $\sqrt{D_i}$), $\sqrt{V_i^N}$, $\sqrt{V_i^{PR}}$, $\sqrt{V_i^{LL}}$ and $\sqrt{V_i^P}$. All the different measures of uncertainty of the empirical Bayes estimates are smaller than the measure of uncertainty of Y_i, the design-based estimates of the median income of the 4-person families. In fact, there is a considerable gain in using empirical Bayes estimator. Generally, both V_i^{PR} and V_i^P are more conservative than V_i^{LL}. It appears that V_i^{PR} is generally slightly more conservative than V_i^P.

2.6. Simulation Experiment

In this section, we conduct a simulation experiment to validate the the proposed measure of uncertainty and also to compare the proposed measure with other rival methods. We consider two values of m, e.g., $m = 20$ and $m = 30$ and consider three different combinations of (σ^2, τ^2) so as to cover all the three cases: (i) $\sigma^2/\tau^2 < 1$. (ii) $\sigma^2/\tau^2 = 1$ and (iii) $\sigma^2/\tau^2 > 1$. We generated 10,000 independent θ_i from a normal distribution with mean zero and variance $\sigma_i^2 = \sigma^2$ and then for each θ_i we generated y_i from a normal with mean θ_i and variance τ^2, $(i = 1, ..., m)$. For each simulation we found the confidence intervals i.e.. $\epsilon_i^{EB} \pm z_{.025}\sqrt{V_i^j}$, where $z_{.025}$ is the upper 2.5% point of standard normal deviate and $j = N, LL, M, PR, Proposed$ and checked whether θ_i belonged to the confidence interval $(i = 1, ..., m)$. We report the average coverage probabilities and the average length for each method on Table 3. In order to investigate the Bayesian coverage, we simulated data only once to find the confidence intervals by various methods. We then generated θ_i 10,000 times from normal with mean is $(1 - B_1)\mu + B_1 y_i$ and the variance is $\tau^2\sigma^2/(\tau^2 + \sigma^2)$ see whether θ_i belonged to the interval $\{\epsilon_i^{EB} \pm z_{.025}\sqrt{V_i^j}\}$. The results are reported in Table 4 and Table 5. In Table 6 we present the relative biases of different MSE estimators which are calculated by 100× [average E(estimator of MSE) - average MSE]/(average of MSE), where the average is over all the small areas and E denotes the simulated average. In Table 7 we present the average simulated MSE of MSE estimators i.e., $E\{V_i^j - MSE_i\}^2$, where V_i^j is the MSE estimator of the ith small area for $j = N, LL, M, PR, Proposed$.

Table 1: Comparison of Different Measures of Uncertainty of Empirical Bayes Estimates for the Baseball Data (Efron and Morris, 1975)

Player Name	Naive	KS I	KS II	Prasad - Rao	Laird - Louis	Proposed
Clemente (Pitts,NL)	.153	.519	.618	.353	.646	.725
F. Robinson (Balt, AL)	.153	.416	.515	.353	.512	.590
F. Howard (Wash, AL)	.153	.327	.427	.353	.396	.474
Johnstone (Cal, AL)	.153	.255	.354	.353	.301	.380
Berry (Chi, AL)	.153	.202	.301	.353	.232	.310
Spencer (Cal, AL)	.153	.202	.301	.353	.232	.310
Kessinger (Chi, NL)	.153	.168	.268	.353	.188	.266
L. Alvarado (Bos, AL)	.153	.154	.253	.353	.169	.248
Santo (Chi, NL)	.153	.162	.261	.353	.179	.258
Swoboda (NY, NL)	.153	.162	.261	.353	.179	.258
Unser (Wash, AL)	.153	.191	.291	.353	.218	.297
Williams (Chi, AL)	.153	.191	.291	.353	.218	.297
Scott (Bos, AL)	.153	.191	.291	.353	.218	.297
Petrocelli (Bos, AL)	.153	.191	.291	.353	.218	.297
E. Rodriguez (KC, AL)	.153	.191	.291	.333	.218	.297
Campaneries (Oak, AL)	.153	.249	.348	.353	.293	.372
Munson (NY, AL)	.153	.332	.432	.353	.402	.481
Alvis (Mill, AL)	.153	.451	.551	.353	.558	.636

Table 2: Comparison of Different Measures of Uncertainty of Empirical Bayes Estimates of Median Incomes of Four-Person Family for 50 States and the District of Columbia for the Year 1988

State No.	$\sqrt{D_i}$	Naive	Prasad - Rao	Laird -Louis	Pro posed	State No.	$\sqrt{D_i}$	Naive	Prasad - Rao	Laird -Louis	Pro posed
1	2183	1265	1284	1266	1276	27	1765	1166	1186	1173	1183
2	3248	1426	1439	1426	1433	28	2632	1335	1352	1336	1344
3	2908	1371	1387	1374	1382	29	3790	1442	1454	1444	1450
4	1989	1246	1266	1248	1258	30	1577	1109	1128	1134	1143
5	3040	1393	1408	1446	1453	31	3094	1380	1395	1381	1388
6	3655	1470	1482	1474	1479	32	3089	1385	1400	1386	1393
7	1972	1231	1250	1232	1242	33	2766	1350	1366	1369	1377
8	1705	1173	1192	1179	1189	34	3006	1368	1384	1371	1378
9	1636	1129	1149	1133	1143	35	2031	1226	1245	1263	1272
10	1576	1108	1127	1108	1118	36	2723	1342	1358	1343	1352
11	3037	1385	1400	1390	1397	37	3122	1381	1395	1383	1390
12	1642	1135	1154	1138	1148	38	2492	1316	1334	1317	1326
13	1744	1166	1185	1167	1176	39	2586	1324	1342	1342	1350
14	1718	1154	1174	1178	1188	40	2867	1356	1372	1356	1364
15	2644	1348	1365	1349	1357	41	2685	1335	1352	1354	1363
16	2169	1260	1279	1264	1274	42	3116	1395	1410	1403	1410
17	3587	1427	1439	1432	1438	43	2733	1340	1357	1340	1348
18	2301	1278	1297	1314	1323	44	3951	1447	1458	1457	1463
19	2329	1281	1300	1317	1326	45	2521	1318	1335	1322	1331
20	2778	1351	1368	1352	1360	46	3019	1387	1402	1395	1403
21	2400	1301	1319	1341	1350	47	2827	1366	1382	1367	1375
22	3393	1425	1438	1435	1442	48	2780	1354	1370	1374	1382
23	3489	1452	1464	1457	1463	49	1731	1164	1183	1173	1182
24	8264	1533	1536	1533	1534	50	3885	1467	1478	1474	1480
25	3106	1403	1418	1407	1415	51	3856	1452	1464	1458	1463
26	1866	1187	1207	1206	1216						

Table 3: Average Frequentist's Coverage Probability and Average Length (nominal coverage =.95)

	$\sigma^2 = 0.5,\ \tau^2 = 0.2$		$\sigma^2 = 1.0,\ \tau^2 = 1.0$		$\sigma^2 = 1.0,\ \tau^2 = 0.5$	
	$m = 20$	$m = 30$	$m = 20$	$m = 30$	$m = 20$	$m = 30$
Naive	0.834	0.845	0.891	0.911	0.845	0.862
	(1.39)	(1.38)	(2.64)	(2.68)	(2.11)	(2.12)
Laird-Louis	0.911	0.900	0.924	0.929	0.911	0.905
	(1.66)	(1.55)	(2.86)	(2.81)	(2.45)	(2.34)
Prasad-Rao	0.937	0.921	0.926	0.929	0.929	0.919
	(1.68)	(1.58)	(2.83)	(2.80)	(2.47)	(2.35)
Proposed	0.946	0.933	0.938	0.939	0.941	0.932
	(1.82)	(1.68)	(2.99)	(2.90)	(2.66)	(2.49)

Table 4: Simulated Bayesian Coverage Probabilities for $m = 20, \sigma^2 = 1.0, \tau^2 = 1.0$

No.	Naive	Laird -Louis	Prasad - Rao	Pro posed
1	0.929	0.939	0.952	0.949
2	0.936	0.938	0.957	0.947
3	0.941	0.943	0.961	0.951
4	0.935	0.940	0.955	0.948
5	0.940	0.944	0.958	0.951
6	0.938	0.941	0.959	0.951
7	0.924	0.946	0.945	0.954
8	0.939	0.941	0.960	0.950
9	0.936	0.940	0.957	0.951
10	0.920	0.952	0.945	0.958
11	0.920	0.953	0.944	0.959
12	0.939	0.942	0.959	0.952
13	0.926	0.949	0.950	0.955
14	0.931	0.946	0.953	0.953
15	0.867	0.960	0.902	0.965
16	0.930	0.944	0.950	0.951
17	0.935	0.937	0.956	0.946
18	0.935	0.946	0.957	0.955
19	0.938	0.939	0.957	0.948
20	0.911	0.951	0.937	0.956

Table 5: Simulated Bayesian Coverage Probabilities for $m = 30, \sigma^2 = 1.0, \tau^2 = 1.0$

No.	Naive	Laird -Louis	Prasad - Rao	Pro posed
1	0.938	0.945	0.954	0.951
2	0.944	0.944	0.960	0.953
3	0.937	0.938	0.953	0.945
4	0.940	0.942	0.956	0.951
5	0.938	0.940	0.955	0.948
6	0.943	0.943	0.961	0.953
7	0.928	0.949	0.948	0.957
8	0.941	0.426	0.956	0.949
9	0.936	0.941	0.953	0.949
10	0.925	0.949	0.944	0.956
11	0.922	0.949	0.940	0.954
12	0.939	0.941	0.955	0.948
13	0.935	0.952	0.953	0.956
14	0.936	0.949	0.955	0.956
15	0.879	0.954	0.907	0.958
16	0.937	0.947	0.955	0.955
17	0.942	0.943	0.950	0.950
18	0.941	0.948	0.957	0.955
19	0.944	0.945	0.959	0.952
20	0.920	0.523	0.943	0.957
21	0.924	0.951	0.944	0.957
22	0.942	0.943	0.956	0.950
23	0.945	0.945	0.959	0.952
24	0.933	0.945	0.951	0.951
25	0.924	0.956	0.947	0.961
26	0.943	0.944	0.958	0.952
27	0.939	0.943	0.954	0.950
28	0.929	0.943	0.948	0.952
29	0.941	0.944	0.957	0.952
30	0.940	0.942	0.957	0.950

Table 6: Percent Average Relative Biases of MSE Estimators

	$\sigma^2 = 0.5,\ \tau^2 = 0.2$		$\sigma^2 = 1.0,\ \tau^2 = 1.0$		$\sigma^2 = 1.0,\ \tau^2 = 0.5$	
	$m = 20$	$m = 30$	$m = 20$	$m = 30$	$m = 20$	$m = 30$
Naive	-22	-19	-17	-12	-18	-22
Laird-Louis	5	-1	-4	-4	-3	1
Prasad-Rao	4	-2	-7	-6	-4	-1
Proposed	23	12	4	2	8	16

Table 7: Average MSE of MSE Estimators (multiply by 10^2)

	$\sigma^2 = 0.5,\ \tau^2 = 0.2$		$\sigma^2 = 1.0,\ \tau^2 = 1.0$		$\sigma^2 = 1.0,\ \tau^2 = 0.5$	
	$m = 20$	$m = 30$	$m = 20$	$m = 30$	$m = 20$	$m = 30$
Naive	0.92	0.74	3.92	2.47	3.06	4.11
Laird-Louis	1.00	0.70	3.13	2.08	2.76	3.92
Prasad-Rao	0.47	0.47	2.25	1.78	2.00	2.04
Proposed	0.98	0.63	2.90	1.87	2.46	3.71

APPENDIX

We shall assume the following regularity conditions throughout the paper. The regularity conditions will be refered to as (RC).

Regularity Conditions (RC):

(a) The elements of X_i and Z_i are uniformly bounded such that $\sum_{i=1}^{m} X_i' V_i^{-1}(\psi) X_i = [O(m)]_{p\times p}$; $i = 1, ..., m$

(b) $\sup_{i\geq 1} n_i < \infty$ and $\sup_{i\geq 1} k_i < \infty$:

(c) $l_i - X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i = [O(1)]_{p\times 1}$;

(d) $\frac{\partial}{\partial \psi_a}[X_i' V_i^{-1}(\psi) Z_i G_i(\psi) \lambda_i] = [O(1)]_{p\times 1}$ for $j = 1, ..., s$

(e) $R_i(\psi) = \sum_{j=0}^{s} \psi_j D_{ij} D_{ij}'$ and $G_i(\psi) = \sum_{j=0}^{s} \psi_j F_{ij} F_{ij}'$, where $\psi_0 = 1$, D_{ij} and F_{ij} $(i = 1, ..., m, j = 0, ..., s)$ are known matrices of order $n_i \times k_i$ and $k_i \times k_i$ respectively and the elements are uniformly bounded known constants such that $R_i(\psi)$ and $G_i(\psi)$ $(i = 1, ..., m)$ are all positive definite matrices. In special cases. some of D_{ij} and F_{ij} may be null matrices.

(f) $\hat{\psi}$ is an estimator of ψ which satisfies (i) $\hat{\psi} - \psi = O_p(m^{-1/2})$, (ii) $\hat{\psi} - \hat{\psi}_{ML} = O_p(m^{-1})$, (iii) $\hat{\psi}(-Y) = \hat{\psi}(Y)$ and (iv) $\hat{\psi}(Y + Xb) = \hat{\psi}(Y)$, for any $b \in R^p$ and for all Y, where $Y = col_{1\leq i\leq m} Y_i$. $X = col_{1\leq i\leq m} X_i$ and $\hat{\psi}_{ML}$ is the maximum likelihood estimator of ψ. Assume that $E(\hat{\psi} - \psi) = -m^{-1}B(\psi) + o(m^{-1})$.

(g) $E(\hat{\beta}(\psi) - \beta)(\hat{\psi} - \psi) = o(m^{-1})$.

Note that conditions (a)-(f) were also needed in Datta and Lahiri (1997) (also see Prasad and Rao 1990). Condition (g) is reasonable following an argument of Cox and Reid (1987).

Let $R_m = O_p(m^{-1})$ and $R_m^* = O_{p^*}(m^{-1})$ denote sequences of random variables such that mR_m and mR_m^* are bounded in probability under Model 1 and Model 2, respectively.

Theorem A.1. Under Model 1 and the regularity conditions (RC),

(i) $E_*[\hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi}^*),\hat{\psi}^*) - \hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi}),\hat{\psi})]^2 = g_{4i}(\hat{\psi};Y_i) + o_p(m^{-1})$;

(ii) $V_i^{LL} = g_{1i}(\hat{\psi}) + g_{2i}(\hat{\psi}) + g_{4i}(\hat{\psi};Y_i) - g_{3i}(\hat{\psi}) - m^{-1}B'(\hat{\psi})\nabla g_{1i}(\hat{\psi}) + o_p(m^{-1})$.

Proof of Theorem A.1.: Using an argument similar to the one given in the proof of Theorem A.1 of Datta and Lahiri (1997), we get

$$\hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi}^*),\hat{\psi}^*) = \hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi}),\hat{\psi}) + (\hat{\psi}^* - \hat{\psi})'\,\nabla\,\hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi})).\hat{\psi}) \\ +O_{p^*}(m^{-1}), \tag{29}$$

where, $\nabla\hat{\theta}(Y_i;\hat{\beta}^*(\hat{\psi})),\hat{\psi}) = \left\{\frac{\partial}{\partial\hat{\psi}_1}\hat{\theta}(Y_i;\hat{\beta}^*(\hat{\psi})),\hat{\psi}),\cdots,\frac{\partial}{\partial\hat{\psi}_s}\hat{\theta}(Y_i;\hat{\beta}^*(\hat{\psi})),\hat{\psi})\right\}'$.
Using $\frac{\partial}{\partial\hat{\psi}_j}\hat{\theta}(Y_i;\hat{\beta}^*(\hat{\psi})),\hat{\psi}) = \sum_{r=1}^{p}\frac{\partial}{\partial\beta_r}\hat{\theta}(Y_i;\beta,\hat{\psi})\mid_{\beta=\hat{\beta}^*(\hat{\psi})} \times \frac{\partial}{\partial\hat{\psi}_j}\hat{\beta}_r^*(\hat{\psi})$
$+ \frac{\partial}{\partial\hat{\psi}_j}\hat{\theta}(Y_i;\beta,\hat{\psi})\mid_{\beta=\hat{\beta}^*(\hat{\psi})}$ and $\frac{\partial}{\partial\hat{\psi}_j}\hat{\beta}_r^*(\hat{\psi}) = O_{p^*}(m^{-1})$ (see Cox and Reid ,1987), we get

$$(\hat{\psi}_j^* - \hat{\psi}_j)\frac{\partial}{\partial\hat{\psi}_j}\hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi}),\hat{\psi}) \\ = (\hat{\psi}_j^* - \hat{\psi}_j)\frac{\partial}{\partial\hat{\psi}_j}\hat{\theta}_i(Y_i;\beta,\hat{\psi})\mid_{\beta=\hat{\beta}^*(\hat{\psi})} + O_{p^*}(m^{-1}). \tag{30}$$

Now,

$$\frac{\partial}{\partial\hat{\psi}_j}\hat{\theta}_i(Y_i;\beta,\hat{\psi})\mid_{\beta=\hat{\beta}^*(\hat{\psi})} \\ = \left\{\frac{\partial}{\partial\hat{\psi}_j}\left[l_i'\beta + \lambda_i' G_i(\hat{\psi})\,Z_i'\,V_i^{-1}(\hat{\psi})\,(Y_i - X_i\beta)\right]\right\}_{\beta=\hat{\beta}^*(\hat{\psi})} \\ = \frac{\partial}{\partial\hat{\psi}_j}\left\{\lambda_i' G_i(\hat{\psi})\,Z_i'\,V_i^{-1}(\hat{\psi})\right\}(Y_i - X_i\hat{\beta}^*(\hat{\psi})) \\ = \frac{\partial}{\partial\hat{\psi}_j}\left\{\lambda_i' G_i(\hat{\psi})\,Z_i'\,V_i^{-1}(\hat{\psi})\right\}\left[Y_i - X_i\hat{\beta}(\hat{\psi}) + X_i\hat{\beta}(\hat{\psi}) - X_i\hat{\beta}^*(\hat{\psi})\right] \\ = L_{ij}'(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi})) + O_{p^*}(m^{-\frac{1}{2}}) \tag{31}$$

where $L_{ij}'(\hat{\psi}) = \frac{\partial}{\partial\hat{\psi}_j}\left\{\lambda_i' G_i(\hat{\psi})\,Z_i'\,V_i^{-1}(\hat{\psi})\right\}$ and $\hat{\beta}^*(\hat{\psi}) - \hat{\beta}(\hat{\psi}) = O_{p^*}(m^{-\frac{1}{2}})$. Using (29) – (31),

$$\hat{\theta}_i(Y_i;\hat{\beta}^*(\hat{\psi}^*).\hat{\psi}^*)$$

$$= \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) + \left\{ \sum_{j=1}^{s} (\hat{\psi}_j^* - \hat{\psi}_j) L_{ij}'(\hat{\psi}) \right\} (Y_i - X_i\hat{\beta}(\hat{\psi})) + O_{p^*}(m^{-1})$$

$$= \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) + (\hat{\psi}^* - \hat{\psi})' L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi})) + O_{p^*}(m^{-1}), \tag{32}$$

where $L_i(\hat{\psi}) = col_{1 \le j \le s} L_{ij}'$.

Using (32), $E_* \, \hat{\beta}^*(\hat{\psi}) = \hat{\beta}(\hat{\psi})$ and $E_* \, (\hat{\psi}^* - \hat{\psi}) = O_p(m^{-1})$, we get

$$\begin{aligned}
& E_* \, \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) \\
= \; & E_* \, \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) + O_p(m^{-1}) \\
= \; & E_* \left\{ l_i' \, \hat{\beta}^*(\hat{\psi}) + \lambda_i' \, G_i(\hat{\psi}) \, Z_i' \, V_i^{-1}(\hat{\psi}) \, (Y_i - X_i\hat{\beta}^*(\hat{\psi})) \right\} + O_p(m^{-1}) \\
= \; & l_i' \, \hat{\beta}(\hat{\psi}) + \lambda_i' \, G_i(\hat{\psi}) \, Z_i' \, V_i^{-1}(\hat{\psi}) \, (Y_i - X_i\hat{\beta}(\hat{\psi})) + O_p(m^{-1}) \\
= \; & \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) + O_p(m^{-1}).
\end{aligned} \tag{33}$$

Now using (33), we get

$$\begin{aligned}
& V_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) \right] \\
= \; & E_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - E_* \{ \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) \} \right]^2 \\
= \; & E_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) + O_p(m^{-1}) \right]^2 \\
= \; & E_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) \right]^2 + O_p(m^{-2}) \\
& + O_p(m^{-1}) E_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) \right] \\
= \; & E_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) \right]^2 + O_p(m^{-2}) \\
= \; & E_* \{ \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) \\
& + \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) \}^2 + O_p(m^{-2}) \\
= \; & E_* Q_1^2 + E_* Q_2^2 + 2 E_* Q_1 Q_2 + O_p(m^{-2}),
\end{aligned} \tag{34}$$

since $E_* \left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi}) \right] = O_p(m^{-1})$, by (33), and using the notation $Q_1 = \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi})$ and $Q_2 = \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi})$. Note that

$$Q_1 = \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) - \hat{\theta}_i(Y_i; \hat{\beta}(\hat{\psi}), \hat{\psi})$$

$$
\begin{aligned}
&= l_i' \hat{\beta}^*(\hat{\psi}) + \lambda_i' G_i(\hat{\psi}) Z_i' V_i^{-1}(\hat{\psi}) (Y_i - X_i\hat{\beta}^*(\hat{\psi})) \\
&\quad -l_i' \hat{\beta}(\hat{\psi}) - \lambda_i' G_i(\hat{\psi}) Z_i' V_i^{-1}(\hat{\psi}) (Y_i - X_i\hat{\beta}(\hat{\psi})) \\
&= l_i' [\hat{\beta}^*(\hat{\psi}) - \hat{\beta}(\hat{\psi})] + \lambda_i' G_i(\hat{\psi}) Z_i' V_i^{-1}(\hat{\psi}) X_i \left\{\hat{\beta}(\hat{\psi}) - \hat{\beta}^*(\hat{\psi})\right\} \\
&= \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\}' \left[\hat{\beta}^*(\hat{\psi}) - \hat{\beta}(\hat{\psi})\right] \qquad (35)
\end{aligned}
$$

and using (32) we get

$$
\begin{aligned}
Q_2 &= \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*) - \hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}), \hat{\psi}) \\
&= (\hat{\psi}^* - \hat{\psi})' L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi})) + O_{p^*}(m^{-1}) \qquad (36)
\end{aligned}
$$

Thus, using (35)

$$
\begin{aligned}
E_* Q_1^2 &= \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\}' E_* \left[\hat{\beta}^*(\hat{\psi}) - \hat{\beta}(\hat{\psi})\right] \\
&\quad \left[\hat{\beta}^*(\hat{\psi}) - \hat{\beta}(\hat{\psi})\right]' \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\} \\
&= \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\}' Var_* [\hat{\beta}^*(\hat{\psi})] \\
&\quad \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\} \\
&= \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\}' \left[\sum_{i=1}^{m} X_i'V_i^{-1}(\hat{\psi})X_i\right]^{-1} \\
&\quad \left\{ l_i - X_i'V_i^{-1}(\hat{\psi}) Z_i G_i(\hat{\psi}) \lambda_i\right\} \\
&= g_{2i}(\hat{\psi}). \qquad (37)
\end{aligned}
$$

Note that in the above $Var_* [\hat{\beta}^*(\hat{\psi})] = \left[\sum_{i=1}^{m} X_i'V_i^{-1}(\hat{\psi})X_i\right]^{-1}$ since $Var[\hat{\beta}(\psi)] = \left[\sum_{i=1}^{m} X_i'V_i^{-1}(\psi)X_i\right]^{-1}$. Using (36)

$$
\begin{aligned}
E_* Q_2^2 &= E_* \left[(\hat{\psi}^* - \hat{\psi})' L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi})) + O_{p^*}(m^{-1})\right]^2 \\
&= E_* \left[(\hat{\psi}^* - \hat{\psi})' L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi}))(Y_i - X_i\hat{\beta}(\hat{\psi}))' L_i'(\hat{\psi})(\hat{\psi}^* - \hat{\psi})\right] \\
&\quad + o_p(m^{-1}) \\
&= E_* \, trace \left[(\hat{\psi}^* - \hat{\psi})' L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi}))(Y_i - X_i\hat{\beta}(\hat{\psi}))'\right. \\
&\quad \left. L_i'(\hat{\psi})(\hat{\psi}^* - \hat{\psi})\right] + o_p(m^{-1})
\end{aligned}
$$

$$
\begin{aligned}
&= trace\ \Big[L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi}))(Y_i - X_i\hat{\beta}(\hat{\psi}))'L_i'(\hat{\psi}) \\
&\quad E_*\ (\hat{\psi}^* - \hat{\psi})(\hat{\psi}^* - \hat{\psi})'\Big] + o_p(m^{-1}) \\
&= trace\ \Big[L_i(\hat{\psi})(Y_i - X_i\hat{\beta}(\hat{\psi}))(Y_i - X_i\hat{\beta}(\hat{\psi}))'L_i'(\hat{\psi})\Sigma(\hat{\psi})\Big] + o_p(m^{-1}) \\
&= g_{4i}(\hat{\psi}; Y_i)\ +\ o_p(m^{-1}).
\end{aligned} \tag{38}
$$

Note that in the above $Var_*(\hat{\psi}^*) = \Sigma(\hat{\psi})$ since $Var(\hat{\psi}) = \Sigma(\psi)$. This proves part (i). Similarly, using (35) $-$ (36)

$$
\begin{aligned}
E_*\, Q_1Q_2 &= \Big[\, l_i\ -\ X_i'V_i^{-1}(\hat{\psi})\ Z_i\ G_i(\hat{\psi})\ \lambda_i\Big]'\ E_*\Big\{(\hat{\beta}^*(\hat{\psi})\ -\ \hat{\beta}(\hat{\psi})) \\
&\quad (\hat{\psi}^*\ -\ \hat{\psi})'\Big\}\ L_i(\hat{\psi})(Y_i\ -\ X_i\ \hat{\beta}(\hat{\psi}))\ +\ o_p(m^{-1}) \\
&= o_p(m^{-1}),
\end{aligned} \tag{39}
$$

Note that in the above $E_*\left[(\hat{\beta}^*(\hat{\psi})\ -\ \hat{\beta}(\hat{\psi}))\ (\hat{\psi}^*\ -\ \hat{\psi})'\right] = o_p(m^{-1})$ by (g) of (RC). Now using (34), and (37) $-$ (39), we get

$$
V_*\left[\hat{\theta}_i(Y_i; \hat{\beta}^*(\hat{\psi}^*), \hat{\psi}^*)\right] = g_{2i}(\hat{\psi})\ +\ g_{4i}(Y_i; \hat{\psi})\ +\ o_p(m^{-1}). \tag{40}
$$

Using Theorem A.2 of Datta and Lahiri (1997), we get

$$
E_*[g_{1i}(\hat{\psi}^*)] = g_{1i}(\hat{\psi}) - m^{-1}B'(\hat{\psi})\nabla g_{1i}(\hat{\psi}) - g_{3i}(\hat{\psi}) + o_p(m^{-1}). \tag{41}
$$

Now, part (ii) follows from (40) and (41).

Theorem A.2. Under Model 1 and the regularity conditions (RC), we have

(i) $E[V_i^P] = r(\hat{\theta}_i^{EB}) + o(m^{-1})$,

(ii) $E[V_i^{LL}] = r(\hat{\theta}_i^{EB}) - 2m^{-1}B'(\hat{\psi})\nabla g_{1i}(\hat{\psi}) - g_{3i}(\hat{\psi}) + o(m^{-1})$.

Proof: Since $\hat{\psi}\ -\ \psi\ =\ o_p(1)$, and $\Sigma(\psi) = O(m^{-1})$ we have $L_i(\hat{\psi})\ =\ L_i(\psi)$ $+\ o_p(1)$, $\hat{\beta}(\hat{\psi}) = \beta + o_p(1)$ and $\Sigma(\hat{\psi}) = \Sigma(\psi) + o_p(m^{-1})$. Thus,

$$
\begin{aligned}
&L_i(\hat{\psi})\left[(Y_i - X_i\hat{\beta}(\hat{\psi}))(Y_i - X_i\hat{\beta}(\hat{\psi}))'\right] L_i'(\hat{\psi})\ \Sigma_i(\hat{\psi}) \\
&= L_i(\psi)\,[(Y_i - X_i\beta)(Y_i - X_i\beta)']\, L_i'(\psi)\ \Sigma_i(\psi) + o_p(m^{-1}).
\end{aligned} \tag{42}
$$

Using (42) and the expressions for $g_{3i}(\psi)$ and $g_{4i}(\hat{\psi}; Y_i)$, we get

$$E[g_{4i}(Y_i; \hat{\psi})] \;=\; g_{3i}(\psi) + o(m^{-1}). \tag{43}$$

Using (13) , (43), $m^{-1}B'(\hat{\psi})\nabla g_{1i}(\hat{\psi}) \;=\; m^{-1}B'(\psi)\nabla g_{1i}(\psi)o_p(m^{-1})$ (which follows from $\hat{\psi} - \psi = o_p(1)$) , $g_{3i}(\hat{\psi}) = g_{3i}(\psi) + o(m^{-1})$ and the expression for V_i^P part (i) of the theorem follows. Now, using part (ii) of Theorem A.1, (13) , (43) and part (i) of this theorem then part (ii) of the theorem follows.

CHAPTER 3

Empirical Bayes Estimation of Small Area Characteristics Under Random Sampling Variances

3.1 Introduction

As explained in Chapter 2, Model 1 covers many small area models considered in the literature. The model assumes a random small area effect through the prior distribution on U_i's, the location parameters. However, random small area effects have not been introduced in the scale parameters. This will result in either very unstable or oversmoothed estimates of the small area variances. To illustrate the point, let θ_i and σ_i^2 be the true mean and true variance of the ith small area respectively ($i = 1, ..., m$). Let y_{ij} be the jth observation in the ith small-area ($i = 1, ..., m; j = 1, ..., n_i$). Conditional on θ_i and σ_i^2, let $y_{i1}, ..., y_{in_i}$ be iid $N(\theta_i, \sigma_i^2)$, $i = 1, ..., m$. Suppose that the primary parameter of interest is σ_i^2 ($i = 1, ..., m$). The assumption that σ_i^2 ($i = 1, ..., m$) are different and fixed parameters will lead to a direct estimator which utilizes the information contained in the observations from the ith small-area alone (e.g.. $S_i^2 = (n_i - 1)^{-1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$, the sample variance of the ith small-area). On the other hand the strong synthetic assumption $\sigma_i^2 = \sigma^2$ ($i = 1, ..., m$) will lead to a oversmooth estimator of σ_i^2 which will use the sample from all areas, e.g., $S^2 = \sum_{i=1}^{m} (n_i - 1) S_i^2 / (n - m)$ (where $n = \sum_{i=1}^{m} n_i$), the pooled sample variance.

The above discussions motivate us to consider a prior distribution for σ_i^2 ($i = 1, ..., m$). Kleffe and Rao (1992) used such a modeling of sampling variances (i.e., σ_i^2) and observed that the formula for the EBLUP of θ_i remains unchanged when compared to the formula given in Prasad and Rao (1990) which assumed $\sigma_i^2 = \sigma^2$ ($i = 1, ..., m$). However, the assumption of random σ_i^2 inflates the mean squared error of EBLUP. Arora and Lahiri (1997) demonstrated that in such a sit-

uation hierarchical Bayes method provides estimates of θ_i which is superior to the corresponding EBLUP. In section 3.2. we state the Bayesian small-area model and discuss the Bayesian estimation of $b(\sigma_i^2)$, a real valued function of σ_i^2. In section 3.3, we obtain the Bayes estimator of θ_i under squared error loss. Replacing prior parameters in the Bayes estimator by their estimates, empirical Bayes estimation of $b(\sigma_i^2)$ and θ_i are considered in section 3.4. A second order Laplace approximation method is proposed in section 3.5 to replace the one-dimensional integrals appearing in the Bayes and EB estimates of $b(\sigma_i^2)$ and θ_i. Finally, in section 3.6. we present results from a simulation.

3.2 The Bayes Estimation of $b(\sigma_i^2)$, a Real Valued Function of σ_i^2

Let y_{ij} denote the value of a characteristic of interest for the jth unit of the ith area ($i = 1, \ldots, m; j = 1, \ldots, n_i$). Let y_i denote the vector of all values from the ith small-area i.e., $y_i = (y_{i1}, \ldots, y_{in_i})$. We shall consider the following model:

MODEL 3

(i) Conditional on θ_i and σ_i^2, y_{ij}'s are independent with
$y_{ij} \mid \theta_i, \sigma_i^2 \sim N(\theta_i, \sigma_i^2)$, $\quad (i = 1, \ldots, m; j = 1, \ldots, n_i)$;

(ii) $\theta_i \overset{ind}{\sim} N(x_i'\beta, \tau^2)$, $\quad (i = 1, \ldots, m)$;

(iii) $\sigma_i^2 \overset{iid}{\sim} IG\{\eta, (\eta - 1)\xi\}$, $\quad (i = 1, \ldots, m)$

where the the density of inverse gamma (IG) is given by $f(\sigma_i^2) = \{(\eta - 1)\xi\}^{\eta} (1/\sigma_i^2)^{\eta+1} e^{-(\eta-1)\xi/\sigma_i^2}/\Gamma(\eta)$, $\sigma_i^2 > 0$. Note that $E(\sigma_i^2) = \xi$ and $var(\sigma_i^2) = \xi^2/(\eta - 2)$, $\eta > 2$. In the above model. we assume that x_i is a $p \times 1$ vector of known and fixed auxiliary variables and β is $p \times 1$ unknown fixed vector.

To get the posterior density of $\sigma_i^2 \mid y_i; \psi$ where $\psi = (\beta, \tau^2, \eta, \xi)$, we first write

down the joint density of $y_i, \theta_i, \sigma_i^2$:

$$\begin{aligned} f(y_i,\theta_i,\sigma_i^2) &\propto (\sigma_i^2)^{-\frac{n_i}{2}} e^{-\frac{1}{2}\sum_{j=1}^{n_i}\frac{1}{\sigma_i^2}(y_{ij}-\theta_i)^2} e^{-\frac{1}{2\tau^2}(\theta_i - x_i'\beta)^2} IG\{\eta,(\eta-1)\xi\} \\ &= c\,(\sigma_i^2)^{-\frac{n_i}{2}} e^{-\frac{1}{2}[\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2+\frac{1}{\tau^2}(\theta_i-x_i'\beta)^2]} IG\{\eta,(\eta-1)\xi\}. \end{aligned} \tag{44}$$

Consider the exponent of (44), i.e.,

$$\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2 + \frac{1}{\tau^2}(\theta_i - x_i'\beta)^2$$

$$\begin{aligned} &= \frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i+\overline{y}_i-\theta_i)^2 + \frac{1}{\tau^2}(\theta_i - x_i'\beta)^2 \\ &= \frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2 + \frac{n_i}{\sigma_i^2}(\overline{y}_i-\theta_i)^2 + \frac{1}{\tau^2}(\theta_i - x_i'\beta)^2, \end{aligned} \tag{45}$$

where $\overline{y}_i = n_i^{-1}\sum_{j=1}^{n_i} y_{ij}$. Now the second and third term of rhs of (45) is

$$\begin{aligned} &= \frac{n_i}{\sigma_i^2}(\overline{y}_i^2+\theta_i^2-2\overline{y}_i\theta_i) + \frac{1}{\tau^2}(\theta_i^2+(x_i'\beta)^2-2\theta_i x_i'\beta) \\ &= (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})\theta_i^2 - 2\theta_i(\frac{n_i\overline{y}_i}{\theta_i}+\frac{x_i'\beta}{\tau^2}) + \frac{n_i\overline{y}_i^2}{\sigma_i^2} + \frac{(x_i'\beta)^2}{\tau^2} \\ &= (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})\left\{\theta_i - (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i\overline{y}_i}{\sigma_i^2}+\frac{x_i'\beta}{\tau^2})\right\}^2 \\ &\quad + \frac{n_i\overline{y}_i^2}{\sigma_i^2} + \frac{(x_i'\beta)^2}{\tau^2} - (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i\overline{y}_i}{\sigma_i^2}+\frac{x_i'\beta}{\tau^2})^2. \end{aligned} \tag{46}$$

Thus, using (46), (45) equates to

$$\begin{aligned} &\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2 + \frac{n_i\overline{y}_i^2}{\sigma_i^2} + \frac{(x_i'\beta)^2}{\tau^2} - (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i\overline{y}_i}{\sigma_i^2}+\frac{x_i'\beta}{\tau^2})^2 \\ &+(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})\left\{\theta_i - (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i\overline{y}_i}{\sigma_i^2}+\frac{x_i'\beta}{\tau^2})\right\}^2. \end{aligned} \tag{47}$$

Replacing (47) into the exponent of (44) and then integrating out θ_i, we get the joint p.d.f of y_i and σ_i^2 as:

$$\begin{aligned} f(y_i,\sigma_i^2) &\propto (\sigma_i^2)^{-\frac{n_i}{2}}(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-\frac{1}{2}} \\ &\quad e^{-\frac{1}{2}\{\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2+\frac{n_i\overline{y}_i^2}{\sigma_i^2}+\frac{(x_i'\beta)^2}{\tau^2}-(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i\overline{y}_i}{\sigma_i^2}+\frac{x_i'\beta}{\tau^2})^2\}} \\ &\quad IG\{\eta,(\eta-1)\xi\}. \end{aligned} \tag{48}$$

Now, let's consider second and third term in the exponent of (48)

$$\frac{n_i\overline{y}_i^2}{\sigma_i^2}+\frac{(x_i'\beta)^2}{\tau^2}-(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i\overline{y}_i}{\sigma_i^2}+\frac{x_i'\beta}{\tau^2})^2$$

$$\begin{aligned}
&= (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})\left\{(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i}{\sigma_i^2}\overline{y}_i^2+\frac{1}{\tau^2}(x_i'\beta)^2)-[(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}(\frac{n_i}{\sigma_i^2}\overline{y}_i+\frac{1}{\tau^2}x_i'\beta)]^2\right\}\\
&= (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})\left\{E(V_i^2)-[E(V_i)]^2\right\}\\
&= (\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})E[V_i-E(V_i)]^2, \qquad (49)
\end{aligned}$$

where the random variable V_i is defined as

$$V_i=\begin{cases}\overline{y}_i & \text{w.p } p_i=(\frac{n_i}{\sigma_i^2}+\frac{1}{\tau^2})^{-1}\frac{n_i}{\sigma_i^2}\\ x_i'\beta & \text{w.p } q_i=1-p_i.\end{cases} \qquad (50)$$

Now

$$\begin{aligned}
E(V_i-E(V_i))^2 &= \left[\overline{y}_i-\overline{y}_ip_i-(1-p_i)x_i'\beta\right]^2p_i\\
&\quad+\left[x_i'\beta-\overline{y}_ip_i-(1-p_i)x_i'\beta\right]^2(1-p_i)\\
&= \left[(1-p_i)\overline{y}_i-(1-p_i)x_i'\beta\right]^2p_i+[-p_i(\overline{y}_i-x_i'\beta)]^2(1-p_i)\\
&= p_i(1-p_i)^2(\overline{y}_i-x_i'\beta)^2+p_i^2(1-p_i)(\overline{y}_i-x_i'\beta)^2\\
&= p_i(1-p_i)(\overline{y}_i-x_i'\beta)^2 \qquad (51)
\end{aligned}$$

Thus, using (51) and some algebra, (49) can be written as

$$\frac{n_i}{n_i\tau^2+\sigma_i^2}(\overline{y}_i-x_i'\beta)^2. \qquad (52)$$

Now, using (48) and (52) the joint distribution of y_i and σ_i^2 can be written as:

$$\begin{aligned}
f(y_i,\sigma_i^2) &\propto (\sigma_i^2)^{-\frac{n_i}{2}}(\frac{n_i\tau^2+\sigma_i^2}{\sigma_i^2\tau^2})^{-\frac{1}{2}}IG\{\eta,\eta(\eta-1)\xi\}\\
&\quad exp\left\{-\frac{1}{2}[\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2+\frac{n_i}{n_i\tau^2+\sigma_i^2}(\overline{y}_i-x_i'\beta)^2]\right\}\\
&= c\,(\sigma_i^2)^{-\frac{1}{2}(n_i-1)}(n_i\tau^2+\sigma_i^2)^{-\frac{1}{2}}IG\{\eta,\eta(\eta-1)\xi\}\\
&\quad exp\left\{-\frac{1}{2}[\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2+\frac{n_i}{n_i\tau^2+\sigma_i^2}(\overline{y}_i-x_i'\beta)^2]\right\}. \qquad (53)
\end{aligned}$$

Finally, the posterior distribution of σ_i^2 is given by

$$
\begin{aligned}
f(\sigma_i^2 \mid y_i;\psi) &\propto (\sigma_i^2)^{-\frac{1}{2}(n_i-1)}(n_i\tau^2+\sigma_i^2)^{-\frac{1}{2}} \\
&\quad exp\left\{-\frac{1}{2}[\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2+\frac{n_i}{n_i\tau^2+\sigma_i^2}(\overline{y}_i-x_i'\beta)^2]\right\} \\
&\quad (\sigma_i^2)^{-\eta-1}exp(-\frac{(\eta-1)\xi}{\sigma_i^2}) \\
&= c\,(\sigma_i^2)^{-(\frac{n_i+1}{2}+\eta)}(n_i\tau^2+\sigma_i^2)^{-\frac{1}{2}} \\
&\quad exp\left\{-\frac{s_i^2}{2\sigma_i^2}-\frac{n_i}{2(n_i\tau^2+\sigma_i^2)}(\overline{y}_i-x_i'\beta)^2-\frac{(\eta-1)\xi}{\sigma_i^2}\right\}, \sigma_i^2>0,
\end{aligned}
\tag{54}
$$

where $s_i^2=\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_i)^2$.

Hence, the Bayes estimator of $b(\sigma_i^2)$ under squared error loss is given by

$$
\hat{b}^B(\sigma_i^2)=E[b(\sigma_i^2)\mid y_i;\psi]=\int_0^\infty b(\sigma_i^2)f(\sigma_i^2\mid y_i;\psi)d\sigma_i^2. \tag{55}
$$

The measure of uncertainty of $\hat{b}^B(\sigma_i^2)$ is measured by $Var[b(\sigma_i^2)\mid y_i;\psi]=E[\{b(\sigma_i^2)\}^2\mid y_i;\psi]-\{E[b(\sigma_i^2)\mid y_i;\psi]\}^2$.

3.3 The Bayes Estimation of θ_i

Using Model 3, we first find the posterior distribution of θ_i given y_i. Now, the joint distribution of y_i, θ_i given σ_i^2 is given by

$$
\begin{aligned}
f(y_i,\theta_i\mid\sigma_i^2) &\propto e^{-\frac{1}{2}\sum_{j=1}^{n_i}\frac{1}{\sigma_i^2}(y_{ij}-\theta_i)^2}e^{-\frac{1}{2\tau^2}(\theta_i-x_i'\beta)^2} \\
&= c\,e^{-\frac{1}{2}[\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2+\frac{1}{\tau^2}(\theta_i-x_i'\beta)^2]},
\end{aligned}
\tag{56}
$$

where c is a constant which does not depend on θ_i. Consider the exponent of (56) i.e.,

$$
\frac{1}{\sigma_i^2}\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2+\frac{1}{\tau^2}(\theta_i-x_i'\beta)^2
$$

$$
=\frac{n_i}{\sigma_i^2}\theta_i^2-2\frac{n_i}{\sigma_i^2}\overline{y}_i\theta_i+\frac{1}{\tau^2}\theta_i^2-2\frac{1}{\tau^2}\theta_i x_i'\beta+c
$$

$$= (\frac{n_i}{\sigma_i^2} + \frac{1}{\tau^2})\theta_i^2 - 2\theta_i(\frac{n_i\bar{y}_i}{\theta_i} + \frac{x_i'\beta}{\tau^2}) + c$$

$$= (\frac{n_i}{\sigma_i^2} + \frac{1}{\tau^2})\left\{\theta_i - (\frac{n_i}{\sigma_i^2} + \frac{1}{\tau^2})^{-1}(\frac{n_i\bar{y}_i}{\sigma_i^2} + \frac{x_i'\beta}{\tau^2})\right\}^2 + c \qquad (57)$$

Hence, the distribution of θ_i given y_i and σ_i^2 is normal with mean $\frac{n_i\tau^2}{\sigma_i^2+n_i\tau^2}\bar{y}_i + \frac{\sigma_i^2}{\sigma_i^2+n_i\tau^2}x_i'\beta$ and variance is $(\frac{n_i}{\sigma_i^2} + \frac{1}{\tau^2})^{-1}$.

Thus, when $\psi = (\beta, \tau^2, \eta, \xi)$ is known. the Bayes estimator of θ_i is given by

$$\begin{aligned}
\hat{\theta}_i^B &= E[\theta_i \mid y_i; \psi] = E\{E[\theta_i \mid y_i, \sigma_i^2; \psi] \mid y_i; \psi\} \\
&= E\left\{\left[\frac{n_i\tau^2}{\sigma_i^2 + n_i\tau^2}\bar{y}_i + \frac{\sigma_i^2}{\sigma_i^2 + n_i\tau^2}x_i'\beta\right] \mid y_i; \psi\right\} \\
&= E\{[(1 - B_i)\bar{y}_i + B_i x_i'\beta] \mid y_i; \psi\} \\
&= (1 - E[B_i \mid y_i; \psi])\bar{y}_i + E[B_i \mid y_i; \psi]x_i'\beta \\
&= (1 - w_i)\bar{y}_i + w_i x_i'\beta \qquad (58)
\end{aligned}$$

where $w_i = E[B_i \mid y_i; \psi] = \int_0^\infty \frac{\sigma_i^2}{\sigma_i^2+n_i\tau^2} f(\sigma_i^2 \mid y_i; \psi)d\sigma_i^2$, $B_i = \sigma_i^2/(\sigma_i^2 + n_i\tau^2)$.

The measure of uncertainty of $\hat{\theta}_i^B$ is given by

$$\begin{aligned}
Var[\theta_i \mid y_i; \psi] &= E\left\{Var[\theta_i \mid y_i, \sigma_i^2; \psi] \mid y_i; \psi\right\} + Var\left\{E[\theta_i \mid y_i, \sigma_i^2; \psi] \mid y_i; \psi\right\} \\
&= E\left\{\frac{\sigma_i^2\tau^2}{\sigma_i^2 + n_i\tau^2} \mid y_i; \psi\right\} + Var\{[(1 - B_i)\bar{y}_i + B_i x_i'\beta] \mid y_i; \psi\} \\
&= \tau^2 w_i + (\bar{y}_i - x_i'\beta)^2 Var[B_i \mid y_i; \psi]. \qquad (59)
\end{aligned}$$

To calculate w_i let us make transformation $B_i = \sigma_i^2/(\sigma_i^2 + n_i\tau^2)$ implies $\sigma_i^2 = n_i\tau^2\frac{B_i}{1-B_i}$ and $\frac{\partial\sigma_i^2}{\partial B_i} = n_i\tau^2(1 - B_i)^{-2}$. Note that the range of B_i is between 0 and 1. Using (54) then the probability distribution of B_i given y_i is

$$\begin{aligned}
\propto\ & B_i^{-(\frac{n_i+1}{2}+\eta)}(1 - B_i)^{-\frac{n_i}{2}-1+\eta} \\
& exp\left\{-\frac{s_i^2(1 - B_i)}{2n_i\tau^2 B_i} - \frac{1 - B_i}{2\tau^2}(\bar{y}_i - x_i'\beta)^2 - \frac{(\eta - 1)\xi(1 - B_i)}{n_i\tau^2 B_i}\right\}. \qquad (60)
\end{aligned}$$

3.4 Empirical Bayes Estimation

In practice, the Bayes estimators $\hat{\theta}_i^B$ and $\hat{b}^B(\sigma_i^2)$ in (55) and (58) involve several unknown parameters $\psi = (\beta, \tau^2, \eta, \xi)$, then we need to estimate them from the

available data. Arora (1994) proposed ANOVA type estimator for each unknown parameter. We will use his estimator to estimate ψ. The estimators of β, ξ, η and τ^2 are given by

$$\hat{\beta} = \left[\sum_{i=1}^{m} x_i'(I_i - n_i^{-1}\hat{\lambda}_i 1_i 1_i')x_i\right]^{-1} \left[\sum_{i=1}^{m} x_i'(I_i - n_i^{-1}\hat{\lambda}_i 1_i 1_i')y_i\right] \tag{61}$$

$$\hat{\xi} = (n_T - m - p)^{-1} \sum_{i=1}^{m} y_i'[I - x_i(x_i x_i')^{-1} x_i] y_i. \tag{62}$$

$$\hat{\eta} = 2 + \hat{\xi}^2/\delta^2 \tag{63}$$

$$\hat{\tau}_*^2 = n_*^{-1} \left\{ \sum_{i=1}^{m} y_i'[I - x_i(x_i x_i')^{-1} x_i] y_i - (n_T - p)\hat{\xi} \right\}, \tag{64}$$

where $\hat{\lambda}_i = n_i\hat{\tau}^2/(\hat{\xi} + n_i\hat{\tau}^2)$, $(i = 1, ...m)$, $n_T = \sum_{i=1}^{m} n_i$, $\delta^2 = \{\sum_{i=1}^{m}(n_i^2 - 1)\}^{-1} \sum_{i=1}^{m} [y_i'[I - x_i(x_i x_i')^{-1} x_i] y_i]^2 - \hat{\xi}^2$. and $n_* = n_T - trace\ [(X'X)^{-1} \sum_{i=1}^{m} n_i^2 \overline{x}_i \overline{x}_i']$. In practice, $\hat{\tau}^2$ might be negative, thus an estimator of τ^2 is $\tau^2 = max(0, \hat{\tau}_*^2)$.

An empirical Bayes estimator of $\hat{\theta}_i^{EB}$ and $\hat{b}^{EB}(\sigma_i^2)$ are now obtained from $\hat{\theta}_i^B$ and $\hat{b}^B(\sigma_i^2)$, respectively when ψ is replaced by $\hat{\psi}$ and are given by

$$\hat{\theta}_i^{EB} = (1 - \hat{w}_i)\overline{y}_i + \hat{w}_i x_i'\hat{\beta} \tag{65}$$

and

$$\hat{b}^{EB}(\sigma_i^2) = E[b(\sigma_i^2) \mid y_i; \hat{\psi}] = \int_0^{\infty} b(\sigma_i^2) f(\sigma_i^2 \mid y_i; \hat{\psi}) d\sigma_i^2. \tag{66}$$

3.5 Second Order Approximation

The empirical Bayes estimator $\hat{\theta}_i^{EB}$ and $\hat{b}^{EB}(\sigma_i^2)$ given in Section 3.4 calculated by numerical integration. We know that posterior probability density is proportional to a likelihood function L and a prior density π. Then the expectation we want to evaluate is in the form

$$E[b(\theta) \mid y] = \frac{\int b(\theta) L(\theta)\pi(\theta) d\theta}{\int L(\theta)\pi(\theta) d\theta} \tag{67}$$

where $b(\theta)$ is a real function of θ (θ is a parameter vector having a posterior density based on a sample n observations). The numerator and denominator of integrals (67) can be evaluated by asymptotic approximation using Laplace's method.

Lemma 1. If h is a smooth function which has first-five derivatives of an m-dimensional vector θ having a minimum at $\hat{\theta}$, and b has first-three derivatives is some other smooth function of θ then, under suitable regularity conditions,

$$E[b(\theta)] \doteq b(\hat{\theta}) + \frac{1}{2n}\left\{b''(\hat{\theta})[h''(\hat{\theta})]^{-1} - b'(\hat{\theta})h'''(\hat{\theta})[h''(\hat{\theta})]^{-2}\right\} \tag{68}$$

see Tierney *et al.* (1989).

Kass and Steffey (1989) pointed out that transforming σ_i^2 to $\sigma_i^2 = exp(-\rho_i)$ to help emphasize that ρ_i is generally preferable to σ_i^2 in numerical work. By using transformation above, we will get the posterior probability density of ρ_i given y_i

$$\begin{aligned} f(\rho_i \mid y_i) \quad &= \quad c\, exp\{\rho_i(\frac{n_i-1}{2}+\eta)\}(n_i\tau^2 + exp(-\rho_i))^{-\frac{1}{2}} \\ & \quad exp\left\{-\frac{\sigma_i^2}{2exp(-\rho_i)} - \frac{n_i}{2(n_i\tau^2+exp(-\rho_i))}(\overline{y}_i - x_i'\beta)^2 - \frac{(\eta-1)\xi}{exp(-\rho_i)}\right\} \end{aligned} \tag{69}$$

To calculate $\hat{w}_i$ i.e., $b(\rho_i) = \frac{exp(-\rho_i)}{n_i\tau^2+exp(-\rho_i)}$, we need $b'(\rho_i)$ and $b''(\rho_i)$ also we should have $h'(\rho_i), h''(\rho_i)$ *and* $h'''(\rho_i)$. where $h(\rho_i) = -n_i^{-1}\log f(\rho_i \mid y_i)$. Taking log on both side of (69), we get

$$\begin{aligned} \log f(\rho_i \mid y_i) \quad &= \quad (\frac{n_i-1}{2}+\eta)\rho_i - \frac{1}{2}\log(n_i\tau^2+exp(-\rho_i)) \\ & \quad -\frac{\sigma_i^2}{2exp(-\rho_i)} - \frac{n_i}{2(n_i\tau^2+exp(-\rho_i))}(\overline{y}_i - x_i'\beta)^2 - \frac{(\eta-1)\xi}{exp(-\rho_i)} + \log c. \end{aligned} \tag{70}$$

The first, second and third derivatives of loglikelihood are

$$\begin{aligned} \frac{\partial \log f(\rho_i \mid y_i)}{\partial \rho_i} \quad &= \quad (\frac{n_i-1}{2}+\eta) + \frac{exp(-\rho_i)}{2(n_i\tau^2+exp(-\rho_i))} - \frac{\sigma_i^2}{2exp(-\rho_i)} \\ & \quad -\frac{n_i(\overline{y}_i - x_i'\beta)^2 exp(-\rho_i)}{2(n_i\tau^2+exp(-\rho_i))^2} - \frac{(\eta-1)\xi}{exp(-\rho_i)}. \end{aligned} \tag{71}$$

$$\frac{\partial^2 \log f(\rho_i \mid y_i)}{\partial \rho_i^2} = -\frac{exp(-\rho_i)}{2(n_i\tau^2 + exp(-\rho_i))} + \frac{exp(-2\rho_i)}{2(n_i\tau^2 + exp(-\rho_i))^2} - \frac{\sigma_i^2}{2exp(-\rho_i)} + \frac{n_i(\overline{y}_i - x_i'\beta)^2 exp(-\rho_i)}{2(n_i\tau^2 + exp(-\rho_i))^2} - \frac{n_i(\overline{y}_i - x_i'\beta)^2 exp(-2\rho_i)}{(n_i\tau^2 + exp(-\rho_i))^3} - \frac{(\eta-1)\xi}{exp(-\rho_i)}, \tag{72}$$

$$\frac{\partial^3 \log f(\rho_i \mid y_i)}{\partial \rho_i^3} = \frac{exp(-\rho_i)}{2(n_i\tau^2 + exp(-\rho_i))} - \frac{3exp(-2\rho_i)}{2(n_i\tau^2 + exp(-\rho_i))^2} + \frac{exp(-3\rho_i)}{(n_i\tau^2 + exp(-\rho_i))^3} - \frac{\sigma_i^2}{2exp(-\rho_i)} - \frac{n_i(\overline{y}_i - x_i'\beta)^2 exp(-\rho_i)}{2(n_i\tau^2 + exp(-\rho_i))^2} + \frac{3n_i(\overline{y}_i - x_i'\beta)^2 exp(-2\rho_i)}{(n_i\tau^2 + exp(-\rho_i))^3} - \frac{3n_i(\overline{y}_i - x_i'\beta)^2 exp(-3\rho_i)}{(n_i\tau^2 + exp(-\rho_i))^4} - \frac{(\eta-1)\xi}{exp(-\rho_i)}, \tag{73}$$

respectively. Noting that $b'(\rho_i) = -\frac{exp(-\rho_i)}{(n_i\tau^2+exp(-\rho_i))} + \frac{exp(-2\rho_i)}{(n_i\tau^2+exp(-\rho_i))^2}$ and $b''(\rho_i) = \frac{exp(-\rho_i)}{(n_i\tau^2+exp(-\rho_i))} - \frac{3exp(-2\rho_i)}{(n_i\tau^2+exp(-\rho_i))^2} + \frac{2exp(-3\rho_i)}{(n_i\tau^2+exp(-\rho_i))^3}$.

An iterative routine is used to get the mode of $\hat{\sigma}_i^2$ or $\hat{\rho}_i$. The first and second derivatives $\nabla f^{(m)}$ and H of function $f(\rho_i \mid y_i)$ are used. In order to solve $f'(\rho_i \mid y_i) = 0$, we employ Newton's iterative procedure.

3.6 Numerical Example

In this section, using a numerical example we will (a) check the accuracy of the Laplace approximation, (b) evaluate the Laird-Louis method of measuring uncertainty of the proposed empirical Bayes estimator and (c) compare the performances of the empirical Bayes estimators of θ_i and σ_i^2.

We generated $m = 30$ independent σ_i^2 $(i = 1, ..., 30)$ from $IG\{\eta, (\eta - 1)\xi\}$ and θ_i $(i = 1, ..., 30)$ from $N(\mu, \tau^2)$. We took $\mu = 10$ and $\tau^2 = 1$ and considered various combinations of (η, ξ). Finally, we generated $y_{ij} \overset{ind}{\sim} N(\theta_i, \sigma_i^2)$ $(i = 1, ...30;\ j = 1, ..., n_i = 10)$.

In Table 8, we present the empirical Bayes estimates of θ_i using numerical integration and Laplace methods. Clearly second order Laplace approximation is

very close to integration, agreeing upto four decimal places most of the time. In Table 9, we tabulate the empirical Bayes estimates $\hat{\sigma}_i^{2\ EB}$, using both numerical integration and Laplace method.

To capture the additional variabilities, Laird and Louis (1987) proposed parametric bootstrap method. Following the Laird-Louis method, we generated $R = 1,000$ the bootstrap samples as follows. For each $r = 1, ..., R$, we generated θ_{ir}^*, where $i = 1, ..., m$, independently from normal with mean $\hat{\mu}$ and variance $\hat{\tau}^2$ and iid σ_{ir}^{2*} from inverse gamma with mean $\hat{\xi}$ and variance $\hat{\xi}^2/(\hat{\eta} - 2)$. We then generated independent bootstrap samples y_{ij}^* from $\sim N(\theta_{ir}^*, \sigma_{ir}^{2*}$. One can simplify the generation by generating $\bar{y}_{ir}^* \overset{ind}{\sim} N(\theta_{ir}^*, \sigma_{ir}^{2*}/n_i)$ and $S_{ir}^2 \overset{ind}{\sim} \sigma_{ir}^{2*}\chi_{n_i-1}^2$. Equation (10) of Laird and Louis (1987) suggests the following measure of variability of e_i^{EB}:

$$Var_i^{EB} = R^{-1} \sum_{r=1}^{R} Var_i^B(y_i; \hat{\psi}_r^*) + (R-1)^{-1} \sum_{r=1}^{R} \{e_i^B(y_i, \hat{\psi}_r^*) - \bar{e}_i^B(y_i)\}^2, \qquad (74)$$

where $\bar{e}_i^B(y_i) = R^{-1} \sum_{r=1}^{R} e_i^B(y_i, \hat{\psi}_r^*)$ and $\hat{\psi}_r^*$ is an estimate of ψ based on the rth bootstrap sample. The Laird-Louis bootstrap measures of uncertainty are report in Table 8. We point out that for 9 of 30 small areas, the Laird-Louis measure is smaller than the Naive measure.

In Table 10-Table 24 , we present the Average Absolute Bias i.e., AAB $= \frac{1}{m}\sum_{i=1}^{m} |T_i - e_i|$, where T_i is θ_i or σ_i^2 and e_i is $\bar{y}, \bar{y}_i, \hat{\theta}_i^{EB}, S^2, S_i^2, \hat{\sigma}_i^{2\ EB}$. Average Squared Deviation (ASD) $= \frac{1}{m}\sum_{i=1}^{m}(T_i - e_i)^2$, Average Relative Bias (ARB) $= \frac{1}{m}\sum_{i=1}^{m} \frac{|T_i - e_i|}{T_i}$ and Average Relative Squared Deviation (ARSD) $= \frac{1}{m}\sum_{i=1}^{m} \frac{(T_i - e_i)^2}{T_i}$.

It turns out that $\hat{\theta}_i^{EB}$ is uniformly better than $\bar{y}_i$ and $\bar{y}$ with respect to all the four measures of evaluation. However, the performance of $\hat{\sigma}_i^{2\ EB}$ depends very much on the value of η. Note that for small η, $Var(\sigma_i^2)$ is large which will support S_i^2 and for large η, $Var(\sigma_i^2)$ is small which will support S^2. Our numerical results agree with the above observation. The empirical Bayes estimator $\hat{\sigma}_i^{2\ EB}$ performs well unless η is very small.

Table 8: Empirical Bayes Estimate $\hat{\theta}_i^{EB}$ of θ_i using Numerical Integration and Laplace methods for $\eta = 10$ and $\xi = 4$

			$\hat{\theta}_i^{EB}$		$Naive(\hat{\theta}_i^{EB})$		$MSE(\hat{\theta}_i^{EB})$
No.	$\overline{y}_i$	θ_i	Integr.	Laplace	Integr.	Laplace	Bootstrap
1	10.52	9.98	10.3705	10.3706	0.2425	0.2422	0.2738
2	10.28	9.19	10.1352	10.1351	0.4289	0.4291	0.2734
3	8.61	10.37	9.1303	9.1305	0.3564	0.3558	0.2981
4	9.76	9.62	9.8292	9.8292	0.3097	0.3097	0.2717
5	10.08	9.79	10.0434	10.0434	0.3005	0.3005	0.2731
6	10.46	10.26	10.3361	10.3362	0.2287	0.2284	0.2730
7	9.32	9.76	9.4991	9.4990	0.2544	0.2541	0.2742
8	9.31	9.71	9.4755	9.4754	0.2299	0.2296	0.2777
9	10.36	9.58	10.2156	10.2155	0.3314	0.3314	0.2720
10	10.57	10.13	10.4239	10.4240	0.2272	0.2269	0.2770
11	9.92	9.99	9.9290	9.9290	0.2158	0.2156	0.2716
12	10.51	10.35	10.3609	10.3610	0.2486	0.2483	0.2728
13	8.82	9.66	9.1299	9.1297	0.2491	0.2483	0.2908
14	7.37	8.70	8.4212	8.4220	0.4074	0.4052	0.3745
15	9.11	8.94	9.3828	9.3828	0.2960	0.2957	0.2842
16	9.79	10.91	9.8413	9.8412	0.2536	0.2535	0.2738
17	8.98	8.68	9.2717	9.2716	0.2711	0.2705	0.2824
18	10.74	10.74	10.4304	10.4303	0.3702	0.3701	0.2780
19	11.84	11.14	11.3218	11.3220	0.2663	0.2641	0.3228
20	11.26	10.92	10.9318	10.9320	0.2374	0.2363	0.2935
21	8.06	8.69	8.5423	8.5420	0.2422	0.2398	0.3260
22	10.07	9.35	10.0365	10.0365	0.2859	0.2859	0.2703
23	10.24	9.81	10.1649	10.1649	0.2609	0.2608	0.2680
24	10.06	10.70	10.0254	10.0254	0.3234	0.3235	0.2716
25	10.33	10.55	10.2447	10.2448	0.2177	0.2174	0.2739
26	9.51	11.13	9.6793	9.6794	0.3440	0.3440	0.2780
27	10.06	9.96	10.0338	10.0338	0.2401	0.2400	0.2703
28	8.61	8.86	8.9738	8.9736	0.2524	0.2512	0.2977
29	13.44	12.67	12.5112	12.5115	0.2958	0.2861	0.4468
30	11.07	10.83	10.7776	10.7778	0.2460	0.2452	0.2912

Table 9: Empirical Bayes Estimate $\hat{\sigma}_i^{2\ EB}$ of σ_i^2 using Numerical Integration and Laplace methods for $\eta = 10$ and $\xi = 4$

			$\hat{\sigma}_i^{2\ EB}$		Laplace	
No.	σ_i^2	S_i^2	Laplace	Integration	Naive MSE	Boots MSE
1	3.737	2.433	3.348	3.374	0.767	3.234
2	6.376	14.182	8.400	8.463	4.779	3.622
3	13.185	7.797	5.782	5.827	2.316	3.339
4	6.335	5.839	4.811	4.848	1.575	3.123
5	2.756	5.350	4.598	4.633	1.439	3.557
6	4.541	1.852	3.093	3.117	0.655	4.188
7	3.693	2.937	3.572	3.599	0.874	5.662
8	3.146	1.866	3.105	3.130	0.661	3.364
9	5.360	7.050	5.340	5.380	1.941	3.473
10	2.683	1.766	3.060	3.084	0.642	3.410
11	2.040	1.364	2.874	2.896	0.565	3.224
12	4.471	2.707	3.467	3.493	0.822	2.904
13	3.305	2.481	3.405	3.432	0.802	3.038
14	7.364	8.336	6.427	6.498	3.025	3.465
15	3.265	4.850	4.420	4.454	1.341	3.152
16	8.775	3.003	3.586	3.613	0.877	4.310
17	4.163	3.557	3.865	3.896	1.029	2.875
18	5.436	9.313	6.350	6.399	2.754	3.285
19	2.960	2.656	3.568	3.598	0.899	3.547
20	2.424	1.908	3.164	3.189	0.694	3.034
21	2.258	1.675	3.124	3.150	0.690	3.997
22	2.989	4.580	4.266	4.299	1.240	3.094
23	5.025	3.331	3.729	3.758	0.949	2.844
24	6.268	6.639	5.154	5.193	1.806	3.356
25	1.724	1.417	2.901	2.923	0.576	3.203
26	5.535	7.796	5.666	5.709	2.185	3.140
27	2.763	2.399	3.323	3.349	0.754	2.819
28	5.481	2.488	3.428	3.456	0.817	2.976
29	3.207	1.585	3.400	3.436	0.889	3.202
30	3.637	2.369	3.352	3.379	0.776	3.766

Table 10: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of θ_i and σ_i^2 where $\eta = 5$ and $\lambda = \xi/\tau^2 = 0.25$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.686	0.132	0.126	0.123	0.099	0.094
ASD	0.767	0.026	0.025	0.024	0.015	0.018
ARB	0.068	0.013	0.013	0.504	0.378	0.356
ARSD	0.007	0.000	0.000	0.413	0.186	0.230

Table 11: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of θ_i and σ_i^2 where $\eta = 5$ and $\lambda = \xi/\tau^2 = 0.50$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.687	0.188	0.171	0.245	0.197	0.187
ASD	0.767	0.052	0.050	0.097	0.061	0.069
ARB	0.069	0.019	0.017	0.504	0.378	0.355
ARSD	0.007	0.000	0.000	0.413	0.186	0.230

Table 12: The Average Absolute Bias. Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of θ_i and σ_i^2 where $\eta = 5$ and $\lambda = \xi/\tau^2 = 1.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.687	0.265	0.234	0.490	0.394	0.374
ASD	0.767	0.104	0.095	0.389	0.246	0.277
ARB	0.069	0.027	0.024	0.504	0.378	0.355
ARSD	0.007	0.001	0.001	0.413	0.186	0.229

Table 13: The Average Absolute Bias. Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of θ_i and σ_i^2 where $\eta = 5$ and $\lambda = \xi/\tau^2 = 2.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.688	0.375	0.319	0.980	0.789	0.747
ASD	0.767	0.207	0.173	1.555	0.983	1.101
ARB	0.069	0.038	0.032	0.504	0.378	0.354
ARSD	0.007	0.002	0.001	0.413	0.186	0.228

Table 14: The Average Absolute Bias. Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of θ_i and σ_i^2 where $\eta = 5$ and $\lambda = \xi/\tau^2 = 4.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.689	0.531	0.419	1.960	1.577	1.493
ASD	0.767	0.415	0.297	6.221	3.931	4.385
ARB	0.069	0.054	0.042	0.504	0.378	0.354
ARSD	0.007	0.004	0.003	0.413	0.186	0.226

Table 15: The Average Absolute Bias. Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 10$ and $\lambda = \xi/\tau^2 = 0.25$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.683	0.140	0.136	0.104	0.114	0.076
ASD	0.767	0.032	0.031	0.021	0.025	0.015
ARB	0.068	0.014	0.014	0.398	0.386	0.255
ARSD	0.007	0.000	0.000	0.250	0.214	0.010

Table 16: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 10$ and $\lambda = \xi/\tau^2 = 0.50$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.682	0.197	0.188	0.209	0.228	0.151
ASD	0.767	0.064	0.060	0.086	0.101	0.059
ARB	0.068	0.020	0.019	0.398	0.386	0.254
ARSD	0.007	0.001	0.001	0.250	0.214	0.010

Table 17: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 10$ and $\lambda = \xi/\tau^2 = 1.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.680	0.279	0.251	0.418	0.457	0.301
ASD	0.767	0.127	0.109	0.343	0.402	0.234
ARB	0.068	0.028	0.025	0.398	0.386	0.252
ARSD	0.007	0.001	0.001	0.250	0.214	0.010

Table 18: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 10$ and $\lambda = \xi/\tau^2 = 2.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.679	0.395	0.322	0.836	0.914	0.599
ASD	0.769	0.254	0.187	1.370	1.608	0.934
ARB	0.067	0.039	0.032	0.398	0.386	0.251
ARSD	0.007	0.003	0.002	0.250	0.214	0.010

Table 19: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 10$ and $\lambda = \xi/\tau^2 = 4.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.679	0.558	0.402	1.671	1.829	1.197
ASD	0.771	0.509	0.293	5.481	6.432	3.703
ARB	0.067	0.056	0.040	0.398	0.386	0.252
ARSD	0.003	0.005	0.003	0.250	0.214	0.010

Table 20: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 100$ and $\lambda = \xi/\tau^2 = 0.25$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.684	0.134	0.132	0.030	0.095	0.028
ASD	0.766	0.025	0.025	0.001	0.015	0.001
ARB	0.068	0.013	0.013	0.117	0.364	0.110
ARSD	0.007	0.000	0.000	0.020	0.202	0.017

Table 21: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 100$ and $\lambda = \xi/\tau^2 = 0.50$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.684	0.189	0.184	0.059	0.191	0.056
ASD	0.766	0.051	0.049	0.005	0.060	0.004
ARB	0.068	0.019	0.018	0.117	0.364	0.110
ARSD	0.007	0.000	0.000	0.020	0.202	0.017

Table 22: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 100$ and $\lambda = \xi/\tau^2 = 1.00$

	θ_i			σ_i^2		
	$\overline{y}$	$\overline{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.684	0.268	0.255	0.118	0.381	0.112
ASD	0.767	0.102	0.094	0.020	0.239	0.017
ARB	0.068	0.027	0.025	0.117	0.364	0.110
ARSD	0.007	0.001	0.001	0.020	0.202	0.017

Table 23: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 100$ and $\lambda = \xi/\tau^2 = 2.00$

	θ_i			σ_i^2		
	$\overline{y}$	$\overline{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.684	0.378	0.340	0.238	0.762	0.223
ASD	0.767	0.203	0.170	0.079	0.956	0.068
ARB	0.068	0.038	0.034	0.117	0.364	0.110
ARSD	0.007	0.002	0.002	0.020	0.202	0.017

Table 24: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates θ_i and σ_i^2 where $\eta = 100$ and $\lambda = \xi/\tau^2 = 4.00$

	θ_i			σ_i^2		
	$\bar{y}$	$\bar{y}_i$	$\hat{\theta}_i^{EB}$	S^2	S_i^2	$\hat{\sigma}_i^{2\ EB}$
AAB	0.683	0.535	0.434	0.476	1.524	0.444
ASD	0.767	0.407	0.288	0.316	3.826	0.272
ARB	0.068	0.053	0.043	0.117	0.364	0.110
ARSD	0.007	0.004	0.003	0.020	0.202	0.017

CHAPTER 4

Empirical Bayes Estimation of Finite Population Means

4.1 Introduction

Suppose a finite population consists of m strata and the i stratum has N_i (known) units. We are interested in estimating certain characteristics for each stratum as well as the entire population. Suppose y be a characteristics of interest and y_{ij} denote the value of the characteristic for the jth element of the ith stratum $(i = 1, ..., m; j = 1, ..., N_i)$. In this chapter, we consider estimation of the mean of each stratum, i.e., $\gamma_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$ $(i = 1, ..., m)$.

The primary source of data is usually available from sample surveys. Suppose a stratified random sample is drawn from the population. Without loss of generality. suppose $y_i = (y_{i1}, ..., y_{in_i})'$ denotes the random sample of size n_i drawn from the ith stratum $(i = 1, ..., m)$. The traditional design unbiased estimator of γ_i is the sample mean given by $\bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$ $(i = 1, ..., m)$.

Ericson (1969a) put forward an elegant formulation of the subjective Bayes approach to the finite population sampling. In his approach, he first assumed that the finite population is a realization from a hypothetical population which is the usual assumption in the super-population approach in finite population theory (see Royall 1970). At the second stage, Ericson (1969a) assumed a subjective prior distribution on the parameters of the super-population model. In practice. it is generally difficult to apply Ericson's Bayesian method since the prior parameters are hardly known. Ghosh and Meeden (1986) considered an empirical Bayes approach under a stratified simple random sampling. using an one-way random effects model. They successfully demonstrated that their method can be very effective in repeated surveys and small-area estimation. Their empirical Bayes estimator is asymptotically optimal in the sense of Robbins (1955). Later on Ghosh and

Lahiri (1987) relaxed the normality assumption of Ghosh and Meeden (1986) and showed that Ghosh-Meeden estimator is robust under the assumption of *posterior linearity* (see Ericson 1969 b; Goldstein 1975; Hartigan 1969). The Ghosh-Meeden empirical Bayes estimator can also be motivated from a best linear unbiased prediction approach of Prasad and Rao (1990). Nandram and Sedransk (1993) extended the Ghosh-Meeden estimator under different but random sampling variances. Recently, Arora *et al.* (1997) considered an alternative to the Nandram-Sedransk method. Their method can incorporate relevant auxiliary information which may be available from various administrative records and censuses. They also proposed, for the first time, a measure of uncertainty of the empirical Bayes estimator of finite population means which can incorporate uncertainty due to estimation of all the parameters in the Bayesian model. Their method is an extension of the parametric bootstrap method proposed earlier by Laird and Louis (1987) to the finite population sampling.

This chapter is a follow up of the method proposed by Arora *et al.* (1997). It is to be noted that the method proposed by Arora *et al.* (1997) involves one-dimensional numerical integration. Although it is not a big problem in calculating the empirical Bayes point estimates, it poses serious problem in finding the measure of uncertainty of the empirical Bayes estimates since several one-dimensional integrals must be calculated at each step of the bootstrap simulation and the accuracy of the numerical integration is hard to check at each step of the Monte Carlo simulation. In order to overcome this difficulty, we propose in this chapter a suitable approximation, using Laplace's second order approximation (see Tierney *et al.* 1989), to the estimation procedure introduced by Arora *et al.* (1997).

In section 4.2, we consider the Bayes and empirical Bayes estimator of the finite population mean γ_i $(i = 1, ..., m)$. In section 4.3, we propose a measure of uncertainty of the Bayes and the empirical Bayes estimators proposed in section 4.2.

4.2 The Bayes and Empirical Bayes Estimation of $\gamma_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$

We shall consider the following model:

MODEL 4

(i) Conditional on θ_i and σ_i^2, y_{ij}'s are independent with

$y_{ij} \mid \theta_i, \sigma_i^2 \sim N(\theta_i, \sigma_i^2)$, $(i = 1, \ldots, m; j = 1, \ldots, N_i)$;

(ii) $\theta_i \overset{ind}{\sim} N(x_i'\beta, \tau^2)$, $(i = 1, \ldots, m)$;

(iii) $\sigma_i^2 \overset{iid}{\sim} IG\{\eta, (\eta-1)\xi\}$. $(i = 1, \ldots, m)$

where the density of inverse gamma (IG) is given by $f(\sigma_i^2) = \{(\eta-1)\xi\}^\eta (1/\sigma_i^2)^{\eta+1} e^{-(\eta-1)\xi/\sigma_i^2}\Gamma(\eta)$, $\sigma_i^2 > 0$. Note that Ghosh-Meeden model can be viewed as a special case of Model 4 when η tends to infinity which implies $\sigma_i^2 = \xi$ $(i = 1, \ldots, m)$. Under the Model 4 and squared error loss i.e., $L(a, \gamma) = m^{-1}\sum_{i=1}^m (a_i - \gamma_i)^2$, the Bayes estimator of γ_i is given by

$$
\begin{aligned}
e_i^{(B)} &= E[\gamma_i \mid y_i; \psi] \\
&= N_i^{-1}\left[\sum_{j=1}^{n_i} y_{ij} + \sum_{j=n_i+1}^{N_i} E(y_{ij} \mid y_i; \psi)\right] \\
&= N_i^{-1}\left[\sum_{j=1}^{n_i} y_{ij} + \sum_{j=n_i+1}^{N_i} E\left\{E(y_{ij} \mid y_i, \theta_i, \sigma_i^2; \psi) \mid y_i; \psi\right\}\right] \\
&= N_i^{-1}\left[\sum_{j=1}^{n_i} y_{ij} + \sum_{j=n_i+1}^{N_i} E(\theta_i \mid y_i; \psi)\right]. \qquad (75)
\end{aligned}
$$

Using (58), (75) follows

$$
\begin{aligned}
e_i^{(B)} &= (1 - f_i)\overline{y}_i + f_i((1 - w_i)\overline{y}_i + w_i x_i'\beta) \\
&= (1 - f_i w_i)\overline{y}_i + f_i w_i x_i'\beta, \qquad (76)
\end{aligned}
$$

where $w_i = E[\sigma_i^2/(\sigma_i^2 + n_i\tau^2) \mid y_i; \psi]$ and $f_i = (N_i - n_i)/N_i$, the finite population correction factor. Empirical Bayes estimation of $\gamma_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$ can be found by replacing w_i with $\hat{w}_i$. Noting that $\hat{w}_i$ will be calculated using procedure in section 3.5.

4.3 Measure of Uncertainty of Empirical Bayes Estimator

It follows from Arora *et al.* (1997) that a measure of variability of empirical Bayes estimator is given by

$$
\begin{aligned}
Var_i^B &= Var_i^B(y_i;\psi) = Var_i^B(\gamma_i \mid y_i;\psi) \\
&= N_i^{-2} Var\left\{ \sum_{j=n_i+1}^{N_i} y_{ij} \mid y_i;\psi \right\} \\
&= N_i^{-2}\left\{ E\left[Var\{ \sum_{j=n_i+1}^{N_i} y_{ij} \mid y_i, \theta_i, \sigma_i^2;\psi\} \mid y_i;\psi \right]\right. \\
&\quad \left. + Var\left[E\{ \sum_{j=n_i+1}^{N_i} y_{ij} \mid y_i, \theta_i, \sigma_i^2;\psi\} \mid y_i;\psi \right]\right\} \\
&= N_i^{-2}[(N_i - n_i)E(\sigma_i^2 \mid y_i;\psi) + (N_i - n_i)^2 Var(\theta_i \mid y_i;\psi)] \\
&= N_i^{-1} f_i E(\sigma_i^2 \mid y_i;\psi) + f_i^2\{Var[E(\theta_i \mid y_i, \sigma_i^2;\psi) \mid y_i;\psi] \\
&\quad + E[Var(\theta_i \mid y_i, \sigma_i^2;\psi) \mid y_i;\psi]\} \\
&= N_i^{-1} f_i E(\sigma_i^2 \mid y_i;\psi) + f_i^2(\overline{y}_i - x_i'\beta)^2 Var(B_i \mid y_i;\psi) \\
&\quad + f_i^2 \tau^2 E(B_i \mid y_i;\psi) \qquad (77)
\end{aligned}
$$

where $B_i = \sigma_i^2/(\sigma_i^2 + n_i\tau^2)$, the expectations and the variance in the last equation of (77) are with respect to the probability distribution of σ_i^2 given y_i and ψ as in (54). We propose that using Laplace method, we calculate $E(Bi \mid y_i;\psi)$ and $Var(B_i \mid y_i;\psi) = E(B_i^2 \mid y_i;\psi) - [E(B_i \mid y_i;\psi)]^2$. We use equation (71) to find $E(Bi \mid y_i;\psi)$ and $E(B_i^2 \mid y_i;\psi)$, where $b(\sigma_i^2) = B_i = [exp(-\rho_i)(exp(-\rho_i + n_i\tau^2)^{-1}]^2$, $b'(\sigma_i^2) = -2b(\sigma_i^2) + 2[b(\sigma_i^2)]^{\frac{3}{2}}$ and $b''(\sigma_i^2) = 4b(\sigma_i^2) - 10[b(\sigma_i^2)]^{\frac{3}{2}} + 6[b(\sigma_i^2)]^2$.

A naive measure of variability of the empirical Bayes estimator e_i^{EB} is obtained as $Var_i^{EB}(y_i;\hat{\psi})$. Note that Var_i^{EB} underestimates the true variability of e_i^{EB} since it does not incorporate the additional variabilities due to estimation of ψ. Prasad and Rao (1990) proposed method based on Taylor series expansion, captures the additional variabilities due to estimation of ψ. Kass and Steffey (1989) proposed a measure of variability for Bayesian approach using Laplace first order approximation. In order to include the additional variability, one can extend the parametric

bootstrap method considered by Laird and Louis (1987) as in (74).

We generated θ_i independently from $N(\mu, \tau^2)$ $(i = 1, ..., 30)$ where $\mu = 10, \tau^2 = 1.0$ and σ_i^2 independently from $\sim IG\{\eta, (\eta-1)\xi\}$ for $\eta = 10, \xi = 4$. We generated $\overline{y}_i \overset{ind}{\sim} N(\theta_i, \sigma_i^2/n_i)$ and $s_i^2 = \sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2 \overset{ind}{\sim} \sigma_i^2\chi^2_{n_i-1}$. We took $n_i = 10$ and $N_i = 100$ for $i = 1, ..., 30$. For bootstrap computation we draw $R = 1,000$ replications for each density above with parameter $\hat{\psi}$. For computation of $\hat{\theta}_i^B(y_i; \hat{\psi}_r^*)$ and $\hat{Var}_i^B(y_i; \hat{\psi}_r^*)$ we considered four cases.

1. When $\hat{\delta} = Var(\sigma_i^2) \leq 0.0$ and $\tau^2 \leq 0.0$, the bootstrap analogue of Model 4 becomes $y_{ij}^* \overset{iid}{\sim} N(\hat{\mu}, \hat{\xi})$. Then

$$\begin{aligned}\hat{\gamma}^B(y_i; \hat{\psi}_r^*) &= (1-f_i)\overline{y}_i + f_i\hat{\mu}_r^* \\ \hat{Var}_i^B(y_i; \hat{\psi}_r^*) &= N_i^{-1}f_i\hat{\xi}_r^*. \end{aligned} \tag{78}$$

2. When $\hat{\delta} > 0.0$ but $\hat{\tau}^2 \leq 0.0$ the bootstrap analogue of Model 4 becomes (i). $y_{ij}^* \mid \hat{\sigma}_i^2 \overset{ind}{\sim} N(\hat{\mu}, \hat{\sigma}_i^{2*})$ and (ii). $\hat{\sigma}_i^{2*} \overset{iid}{\sim} IG(\hat{\eta}, (\hat{\eta}-1)\hat{\xi})$. Then the bootstrap empirical Bayes is

$$\begin{aligned}\hat{\gamma}^B(y_i; \hat{\psi}_r^*) &= (1-f_i)\overline{y}_i + f_i\hat{\mu}_r^* \\ \hat{Var}_i^B(y_i; \hat{\psi}_r^*) &= N_i^{-1}f_iE_*[\sigma_i^{2*} \mid y_i^*; \hat{\psi}_r^*]. \end{aligned} \tag{79}$$

where $E_*[\sigma_i^{2*} \mid y_i; \hat{\psi}_r^*] = \frac{n_i(\overline{y}_i^* - \hat{\mu}^*)^2 + s_i^{2*} + 2\hat{\xi}^*(\hat{\eta}^*-1)}{n_i + 2\hat{\eta}^* - 2}$. From (i). and (ii). we have the density $f(\sigma_i^{2*} \mid y_i^*; \hat{\psi}) \propto \sigma_i^{2*-(n_i/2+\hat{\eta}-1)} exp - \frac{1}{2\sigma_i^{2*}}\{n_i(\overline{y}_i^* - \hat{\mu})^2 + s_i^{2*} + 2\hat{\xi}(\hat{\eta}-1)\}$. Now $E_*[\sigma_i^{2*} \mid y_i^*; \hat{\psi}]$ can be easily found.

3. When $\hat{\delta} \leq 0.0$ but $\hat{\tau}^2 > 0.0$ the bootstrap analogue of Model 4 becomes (i). $y_{ij}^* \mid \theta_i^*; \hat{\xi} \overset{ind}{\sim} N(\theta_i^*, \hat{\xi})$ and (ii). $\theta_i^* \overset{iid}{\sim} N(\hat{\mu}, \hat{\tau}^2)$. Then the bootstrap empirical Bayes is

$$\begin{aligned}\hat{\gamma}^B(y_i; \hat{\psi}_r^*) &= (1 - f_i\hat{w}_i^*)\overline{y}_i + f_i\hat{w}_i^*\hat{\mu}_r^* \\ \hat{Var}_i^{EB}(\hat{\psi}_r^*) &= N_i^{-1}f_i\hat{\xi}^* + f_i^2\hat{\tau}^{2*}\hat{\xi}^*/(\hat{\xi}^* + n_i\hat{\tau}^{2*}), \end{aligned} \tag{80}$$

where $\hat{w}_i^* = \hat{\tau}^{2*}\hat{\xi}^*/(\hat{\xi}^* + n_i\hat{\tau}^{2*})$.

4. When $\hat{\delta} > 0.0$ and $\hat{\tau}^2 > 0.0$ then we use Model 4 with the usual and using

equation (76) and (77). In Table 25 we present the results of the measure of uncertainty of bootstrap method as well as naive method. In Table 25 we report the AAB, ASD, ARB and ARSD of different η and $\lambda = \xi/\tau^2$. It turns out that the proposed estimator is uniformly better than the survey estimator.

Table 25: Empirical Bayes Estimate e_i^{EB} of γ_i where $\eta = 10$ and $\xi = 4$

					$MSE(e_i^{EB})$	
No.	γ_i	$\overline{y}$	$\overline{y}_i$	e_i^{EB}	Naive	Bootstrap
1	10.204	10.251	10.408	10.350	0.202	0.210
2	9.017	10.251	9.296	9.698	0.238	0.231
3	9.780	10.251	11.141	10.605	0.375	0.231
4	9.857	10.251	9.820	10.007	0.243	0.213
5	9.943	10.251	10.204	10.222	0.203	0.209
6	10.425	10.251	10.842	10.635	0.191	0.218
7	9.929	10.251	11.098	10.749	0.231	0.228
8	9.510	10.251	9.525	9.793	0.203	0.221
9	9.234	10.251	10.377	10.325	0.227	0.208
10	9.966	10.251	9.901	10.021	0.185	0.210
11	9.903	10.251	10.417	10.367	0.160	0.208
12	10.169	10.251	10.451	10.378	0.198	0.212
13	10.092	10.251	9.791	9.990	0.241	0.213
14	8.648	10.251	9.280	9.778	0.301	0.234
15	9.037	10.251	9.652	9.856	0.185	0.217
16	10.671	10.251	9.813	9.992	0.226	0.214
17	8.590	10.251	7.314	8.734	0.326	0.447
18	10.323	10.251	10.367	10.322	0.214	0.207
19	11.173	10.251	11.501	11.089	0.184	0.254
20	10.923	10.251	10.772	10.564	0.221	0.215
21	8.601	10.251	8.994	9.424	0.191	0.254
22	9.553	10.251	9.038	9.594	0.266	0.248
23	9.825	10.251	11.038	10.761	0.193	0.226
24	10.829	10.251	11.307	10.826	0.261	0.242
25	10.418	10.251	10.685	10.541	0.180	0.213
26	10.802	10.251	10.643	10.510	0.184	0.213
27	9.844	10.251	9.930	10.037	0.179	0.210
28	8.960	10.251	9.900	10.054	0.246	0.212
29	12.627	10.251	11.710	11.254	0.175	0.268
30	11.062	10.251	12.313	11.528	0.229	0.331

Table 26: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of γ_i where $\eta = 5$

	$\lambda = 0.25$			$\lambda = 0.50$		
	$\bar{y}$	$\bar{y}_i$	e_i^{EB}	$\bar{y}$	$\bar{y}_i$	e_i^{EB}
AAB	0.6851	0.1497	0.1470	0.6878	0.2118	0.2029
ASD	0.7762	0.0338	0.0325	0.7858	0.0676	0.0632
ARB	0.0689	0.0150	0.0148	0.0694	0.0212	0.0206
ARSD	0.0077	0.0003	0.0003	0.0079	0.0007	0.0007

	$\lambda = 1.00$			$\lambda = 2.00$		
	$\bar{y}$	$\bar{y}_i$	e_i^{EB}	$\bar{y}$	$\bar{y}_i$	e_i^{EB}
AAB	0.6982	0.2995	0.2797	0.7158	0.4235	0.3836
ASD	0.8052	0.1351	0.1204	0.8440	0.2703	0.2216
ARB	0.0708	0.0300	0.0284	0.0731	0.0424	0.0390
ARSD	0.0082	0.0014	0.0013	0.0088	0.0027	0.0024

	$\lambda = 4.00$		
	$\bar{y}$	$\bar{y}_i$	e_i^{EB}
AAB	0.7590	0.5990	0.5212
ASD	0.9214	0.5406	0.3878
ARB	0.0781	0.0600	0.0533
ARSD	0.0099	0.0054	0.0042

Table 27: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of γ_i where $\eta = 10$

	$\lambda = 0.25$			$\lambda = 0.50$		
	$\bar{y}$	$\bar{y}_i$	e_i^{EB}	$\bar{y}$	$\bar{y}_i$	e_i^{EB}
AAB	0.6680	0.1252	0.1222	0.6759	0.2020	0.1997
ASD	0.7532	0.0230	0.0224	0.7593	0.0665	0.0606
ARB	0.0667	0.0125	0.0122	0.0682	0.0203	0.0203
ARSD	0.0073	0.0002	0.0002	0.0076	0.0007	0.0006

	$\lambda = 1.00$			$\lambda = 2.00$		
	$\bar{y}$	$\bar{y}_i$	e_i^{EB}	$\bar{y}$	$\bar{y}_i$	e_i^{EB}
AAB	0.6768	0.2695	0.2656	0.6855	0.3811	0.3535
ASD	0.7621	0.1176	0.1125	0.7768	0.2352	0.2093
ARB	0.0685	0.0273	0.0270	0.0698	0.0387	0.0360
ARSD	0.0077	0.0012	0.0012	0.0081	0.0025	0.0022

	$\lambda = 4.00$		
	$\bar{y}$	$\bar{y}_i$	e_i^{EB}
AAB	0.7033	0.5389	0.4607
ASD	0.8141	0.4703	0.3677
ARB	0.0721	0.0548	0.0471
ARSD	0.0087	0.0050	0.0039

Table 28: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of γ_i where $\eta = 100$

	$\lambda = 0.25$			$\lambda = 0.50$		
	$\overline{y}$	$\overline{y}_i$	e_i^{EB}	$\overline{y}$	$\overline{y}_i$	e_i^{EB}
AAB	0.6669	0.1415	0.1344	0.6610	0.2002	0.1866
ASD	0.7561	0.0284	0.0276	0.7538	0.0567	0.0534
ARB	0.0664	0.0141	0.0134	0.0657	0.0200	0.0185
ARSD	0.0072	0.0003	0.0003	0.0071	0.0006	0.0005

	$\lambda = 1.00$			$\lambda = 2.00$		
	$\overline{y}$	$\overline{y}_i$	e_i^{EB}	$\overline{y}$	$\overline{y}_i$	e_i^{EB}
AAB	0.6900	0.2641	0.2619	0.7059	0.3735	0.3582
ASD	0.7805	0.1149	0.1097	0.8104	0.2298	0.2061
ARB	0.0710	0.0263	0.0262	0.0722	0.0372	0.0360
ARSD	0.0080	0.0011	0.0011	0.0086	0.0023	0.0021

	$\lambda = 4.00$		
	$\overline{y}$	$\overline{y}_i$	e_i^{EB}
AAB	0.7354	0.5282	0.4908
ASD	0.8769	0.4596	0.3685
ARB	0.0758	0.0526	0.0498
ARSD	0.0095	0.0046	0.0038

CHAPTER 5

Empirical Bayes Estimation of Finite Population Variances

5.1 Introduction

In this chapter, we consider the setting of Chapter 4 but consider the estimation of the strata variances, i.e., $\gamma_i = N_i^{-1}\sum_{j=1}^{N_i}(y_{ij}-\mu_i)^2$, where μ_i denotes the mean of the ith stratum $(i = 1, ..., m)$.

For the last fifteen years, there has been a growing demand from both the public and private sectors to produce reliable statistics for various subgroups of a finite population. According to Brakstone (1987) "there is in Canada, and probably in other countries too, an increasing government concern with issues of distribution , equity and disparity." Consider the problem of comparing the income distribution for various geographical areas of a country. Is it enough to consider just the per-capita income? Probably not, since two geographical areas may be comparable in terms of their per-capita incomes, yet they may vary considerably in terms of diversity which can be measured by the variances of their income distributions. Although the problem of finite population variances for different geographic groups is a very important problem, it has received relatively less attention than the problem of estimation of means, ratios and proportions for different geographical areas in the finite population sampling.

Ericson (1969a) briefly addressed the problem of the Bayesian estimation of a finite population variance under simple random sampling. Datta and Ghosh (1993) provided a unified approach to the Bayesian estimation of different strata variances in finite population sampling under stratified random sampling. Ghosh and Lahiri (1987) considered the problem using a linear empirical Bayes approach. Lahiri and Tiwari (1990) proposed a nonparametric empirical Bayes estimation using the Dirichelet process prior (Ferguson 1973).

Note that the model considered by Datta and Ghosh (1993) does not incorporate stratum specific random effects through the scale components. Although, this synthetic assumption may have insignificant effect in the estimation of different stratum means, it may cause unduly shrinkage in the Bayes estimator of different stratum variances. Ghosh and Lahiri (1987) and Lahiri and Tiwari (1990) introduced random stratum effects through the scale parameters, but even then they failed to overcome the overshrinkage problem primary because of the linear nature of their Bayes estimators. However, we realize that the linear empirical Bayes procedure of Ghosh and Lahiri (1987) and the nonparametric empirical Bayes approach of Lahiri and Tiwari (1990) are very robust and it is difficult to resolve the problem associated with overshrinking without being specific about the distribution of the stratum specific random scale effects.

In section 5.2, we propose the Bayes and empirical Bayes estimator of the ith stratum variance. In Section 5.3, we present a measure of uncertainty of the estimators proposed in section 5.2.

5.2 The Bayes and Empirical Bayes Estimation of the Strata Variances

Recall the Model introduced in chapter 4 :

MODEL 4

(i) Conditional on θ_i and σ_i^2, y_{ij}'s are independent with
$y_{ij} \mid \theta_i, \sigma_i \sim N(\theta_i, \sigma_i^2)$, $(i = 1, \ldots, m; j = 1, \ldots, N_i)$;

(ii) $\theta_i \overset{ind}{\sim} N(x_i'\beta, \tau^2)$, $\qquad (i = 1, \ldots, m)$;

(iii) $\sigma_i^2 \overset{ind}{\sim} IG\{\eta, (\eta - 1)\xi\}$ $\quad (i = 1, \ldots, m)$

where the density of inverse gamma (IG) has mean ξ and variance $\xi^2/(\eta - 2)$.

Suppose a sample of size n_i is collected from ith finite population, i.e., we have

the observation vector $y_i = (y_{i1}, \ldots, y_{in_i})$. The Bayes estimator of $\gamma(y_i)$, under squared error loss, is

$$e_i^B = E(\gamma_i(y_i) \mid y_i). \tag{S1}$$

Let $s_i = \{1, \ldots, n_i\}$ be the set of labels of units in the sample.

The goal now is to estimate γ_i, the variance for the ith finite population (small area). To get the Bayes estimator $\gamma_i(y_i) = N_i^{-1} \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2$, Arora (1994) proved the following lemma.

Lemma 2 : For any N real numbers, a_i, $(i = 1, \ldots, N)$,

$$\sum_{i \neq j}^{N} (a_i - a_j)^2 = 2N \sum_{i=1}^{N} (a_i - \bar{a})^2, \ \textit{where } \bar{a} = N^{-1} \sum_{i=1}^{N} a_i \ . \tag{S2}$$

Using Lemma 2, γ_i can be written as

$$\gamma_i = \frac{1}{2} N_i^{-2} \sum_{1 \leq j \neq j' \leq N_i} (y_{ij} - y_{ij'})^2. \tag{S3}$$

Three cases are possible:

$$\begin{aligned} &(i) \quad j \in s_i \textit{ and } j' \in s_i, \\ &(ii) \quad j \in s_i \textit{ but } j' \notin s_i, \\ &(iii) \quad j \notin s_i \textit{ and } j' \notin s_i. \end{aligned}$$

Equation (S3) can be written as

$$\begin{aligned} \gamma_i = \frac{1}{2} N_i^{-2} \Bigg[& \sum_{1 \leq j \neq j' \leq n_i} (y_{ij} - y_{ij'})^2 \\ & + 2 \sum_{j=1}^{n_i} \sum_{j'=n_i+1}^{N_i} (y_{ij} - y_{ij'})^2 \\ & + \sum_{n_i+1 \leq j \neq j' \leq N_i} (y_{ij} - y_{ij'})^2 \Bigg]. \end{aligned} \tag{S4}$$

Using (84) in (81), one gets,

$$
\begin{aligned}
e_i^B &= E[\gamma_i \mid y_i] \\
&= \frac{1}{2} N_i^{-2} \Bigg[\sum_{1 \le j \ne j' \le n_i} (y_{ij} - y_{ij'})^2 \\
&\quad + 2 \sum_{j=1}^{n_i} \sum_{j'=n_i+1}^{N_i} E[(y_{ij} - y_{ij'})^2 \mid y_i] \\
&\quad + \sum_{n_i+1 \le j \ne j' \le N_i} E[(y_{ij} - y_{ij'})^2 \mid y_i] \Bigg]. \qquad (85)
\end{aligned}
$$

Expectation in the second term of (85) can be written as

$$
\begin{aligned}
& E[\{(y_{ij} - \theta_i) - (y_{ij'} - \theta_i)\}^2 \mid y_i] \\
= \; & E[(y_{ij} - \theta_i)^2 \mid y_i] + E[(y_{ij'} - \theta_i)^2 \mid y_i] \\
& -2E[(y_{ij} - \theta_i)(y_{ij'} - \theta_i) \mid y_i]. \qquad (86)
\end{aligned}
$$

First term in (86) i.e.,

$$E[(y_{ij} - \theta_i)^2 \mid y_i]$$

$$
\begin{aligned}
&= E[\{(y_{ij} - E(\theta_i \mid y_i)) - (\theta_i - E(\theta_i \mid y_i))\}^2 \mid y_i] \\
&= \{y_{ij} - E(\theta_i \mid y_i)\}^2 + E[\{\theta_i - E(\theta_i \mid y_i)\}^2 \mid y_i] \\
&\quad -2E[\{y_{ij} - E(\theta_i \mid y_i)\}\{\theta_i - E(\theta_i \mid y_i)\} \mid y_i] \\
&= [y_{ij} - E(\theta_i \mid y_i)]^2 + Var(\theta_i \mid y_i). \qquad (87)
\end{aligned}
$$

It was shown in Chapter 4 (Mean estimation problem), that $E(\theta_i \mid y_i) = (1 - w_i)\overline{y}_i + w_i x_i'\beta$, where $w_i = E[b(\sigma_i^2) \mid y_i]$, with $b(\sigma_i^2) = (1 + n_i\tau^2\sigma_i^2)^{-1}$. E represents the expectation using Laplace method. Thus, the first term of (86) becomes

$$
\begin{aligned}
&= Var(\theta_i \mid y_i) + [y_{ij} - (1 - w_i)\overline{y}_i - w_i x_i'\beta]^2 \\
&= Var(\theta_i \mid y_i) + [y_{ij} - \overline{y}_i + w_i(\overline{y}_i - x_i'\beta)]^2 \\
&= Var(\theta_i \mid y_i) + (y_{ij} - \overline{y}_i)^2 + w_i^2(\overline{y}_i - x_i'\beta)^2 \\
&\quad +2(y_{ij} - \overline{y}_i)w_i(\overline{y}_i - x_i'\beta) \qquad (88)
\end{aligned}
$$

Second term of (86)

$$
\begin{aligned}
& E[(y_{ij'} - \theta_i)^2 \mid y_i] \\
= \;& E\left[E\{(y_{ij'} - \theta_i)^2 \mid y_i, \sigma_i^2, \theta, \} \mid y_i\right] \\
= \;& E(\sigma_i^2 \mid y_i). \qquad (89)
\end{aligned}
$$

Third term of (86)

$$
\begin{aligned}
= \;& E[(y_{ij})(y_{ij'} - \theta_i) \mid y_i] - E[\theta_i(y_{ij'} - \theta_i) \mid y_i] \\
= \;& E\left[(y_{ij})E\{(y_{ij'} - \theta_i) \mid y_i, \theta, \sigma_i^2\} \mid y_i\right] \\
& -E\left[E\{\theta_i(y_{ij'} - \theta_i) \mid y_i, \theta, \sigma_i^2\} \mid y_i\right] \\
= \;& 0. \qquad (90)
\end{aligned}
$$

Using (88), (89), and (90) in (86), second term of (85), where $j \in s_i$, $j' \notin s_i$. i.e..

$$
E[(y_{ij} - y_{ij'})^2 \mid y_i]
$$

$$
\begin{aligned}
= \;& Var(\theta_i \mid y_i) + E(\sigma_i^2 \mid y_i) + (y_{ij} - \overline{y}_i)^2 \\
& +w_i^2(\overline{y}_i - x_i'\beta)^2 + 2w_i(y_{ij} - \overline{y}_i)(\overline{y}_i - x_i'\beta). \qquad (91)
\end{aligned}
$$

Now, the third expectation term of (85) is $E[(y_{ij} - y_{ij'})^2 \mid y_i]$, where $j, j' \notin s_i$

$$
\begin{aligned}
= \;& E[\{(y_{ij} - \theta_i) - (y_{ij'} - \theta_i)\}^2 \mid y_i] \\
= \;& E[(y_{ij} - \theta_i)^2 \mid y_i] + E[(y_{ij'} - \theta_i)^2 \mid y_i] \\
& -2E[(y_{ij} - \theta_i)(y_{ij'} - \theta_i) \mid y_i] \qquad (92)
\end{aligned}
$$

First term in (92)

$$
\begin{aligned}
= \;& E\left[E\{(y_{ij} - \theta_i)^2 \mid y_i, \sigma_i^2, \theta\} \mid y_i\right] \\
= \;& E(\sigma_i^2 \mid y_i). \qquad (93)
\end{aligned}
$$

Similarly second term in (92)

$$
E(\sigma_i^2 \mid y_i). \qquad (94)
$$

The third term of (92) is

$$
\begin{aligned}
&= E[(y_{ij})(y_{ij'} - \theta_i) \mid y_i] - E[\theta_i(y_{ij'} - \theta_i) \mid y_i] \\
&= E\left[(y_{ij})E\{y_{ij'} - \theta_i) \mid y_i, \theta, \sigma_i^2\} \mid y_i\right] \\
&\quad -E\left[E\{\theta_i(y_{ij'} - \theta_i) \mid y_i, \theta, \sigma_i^2\} \mid y_i\right] \\
&= 0. \qquad (95)
\end{aligned}
$$

Combining (93), (94) and (95) the third term of (86) becomes

$$
2E(\sigma_i^2 \mid y_i) \qquad (96)
$$

Now, using (91) and (96) in (95), the Bayes estimator of γ_i

$$
\begin{aligned}
e_i^B &= \frac{1}{2}N_i^{-2}\Bigg[\sum_{1\le j\ne j'\le n_i}(y_{ij} - y_{ij'})^2 \\
&\quad +2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}\Big\{Var(\theta_i \mid y_i) + E(\sigma_i^2 \mid y_i) \\
&\quad +(y_{ij} - \overline{y}_i)^2 + w_i^2(\overline{y}_i - x_i'\beta)^2 \\
&\quad + 2w_i(y_{ij} - \overline{y}_i)(\overline{y}_i - x_i'\beta)\Big\} \\
&\quad + 2\sum_{n_i+1\le j\ne j'\le N_i} E(\sigma_i^2 \mid y_i)\Bigg]. \qquad (97)
\end{aligned}
$$

Using $s_i^2 = (n_i - 1)^{-1}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2$, one gets from (97)

$$
\begin{aligned}
e_i^B &= N_i^{-2}\Big[n_i(n_i - 1)s_i^2 + (N_i - n_i)n_iVar(\theta_i \mid y_i) \\
&\quad (N_i - n_i)n_iE(\sigma_i^2 \mid y_i) + (n_i - 1)(N_i - n_i)s_i^2 \\
&\quad +(N_i - n_i)n_iw_i^2(\overline{y}_i - x_i'\beta)^2 \\
&\quad +2(N_i - n_i)w_i\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)(\overline{y}_i - x_i'\beta) \\
&\quad (N_i - n_i)(N_i - n_i - 1)E(\sigma_i^2 \mid y_i)\Big] \\
&= N_i^{-2}\Big[N_i(n_i - 1)s_i^2 + (N_i - n_i)n_i\{Var(\theta_i \mid y_i) \\
&\quad +w_i^2(\overline{y}_i - x_i'\beta)^2\} + (N_i - n_i)(N_i - 1)E(\sigma_i^2 \mid y_i) \qquad (98)
\end{aligned}
$$

Using the notations of Chapter 4, $f_i = (N_i - n_i)/N_i$, the Bayes estimator of γ_i can be written as

$$
e_i^B = N_i^{-1}(n_i - 1)s_i^2 + f_i(1 - f_i)\{Var(\theta_i \mid y_i) + w_i^2(\overline{y}_i - x_i'\beta)^2\}
$$

$$+\frac{N_i - 1}{N_i} f_i E(\sigma_i^2 \mid y_i) \tag{99}$$

The empirical Bayes Estimator for $\hat{\gamma}_i$ under quared error loss function can be found by replacing ψ with $\hat{\psi}$ in equation (99).

5.3 Measure of Uncertainty of Strata Variance Estimator

Under the Model 4 , variance estimator of $N_i^{-1} \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2$ is given by

$$\begin{aligned} & Var\left\{ N_i^{-1} \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2 \mid y_i \right\} \\ = \; & N_i^{-2} \left(Var\left[E\left\{ \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi \right\} \mid y_i \right] \right. \\ & \left. + E\left[Var\left\{ \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi \right\} \mid y_i \right] \right). \end{aligned} \tag{100}$$

The first term of expectation of (100) follows from derivation of Section 5.2, after some algebra, however the super population model of y_{ij} given all others are known i.e.,

$$\begin{aligned} & E\left\{ \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi \right\} \\ = \; & N_i^{-1} \{ n_i(n_i - 1) s_i^2 + (N_i - n_i)(N_i + n_i - 1)\sigma_i^2 \}. \end{aligned} \tag{101}$$

Now, using (101) the first term of (100) becomes

$$f_i^2 (N_i + n_i - 1)^2 Var(\sigma_i^2 \mid y_i) \tag{102}$$

The second term of variance part of (100)

$$\begin{aligned} & Var\left\{ \sum_{j=1}^{N_i} (y_{ij} - \overline{y}_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi \right\} \\ = \; & Var\left[\frac{1}{2} N_i^{-1} \left\{ \sum_{1 \leq j \neq j' \leq n_i} [(y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi] \right. \right. \end{aligned}$$

$$+2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}[(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i.\sigma_i^2;\psi]$$

$$\left.\left.+\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}\right]$$

$$= \frac{1}{4}N_i^{-2}\left[Var\left\{2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}[(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi]\right.\right.$$

$$\left.\left.+\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}\right]$$

$$= \frac{1}{4}N_i^{-2}\left[Var\left\{2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}\right.$$

$$+Var\left\{\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}$$

$$+2Cov\left\{2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\},$$

$$\left.\left\{\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i.\sigma_i^2;\psi\right\}\right]. \qquad (103)$$

The first term of (103)

$$Var\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}$$

$$= E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2$$

$$-\left\{E\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2. \qquad (104)$$

The first term of (104) i.e.,

$$= E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2$$

$$= E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}\left[(y_{ij}-\theta_i)^2+(y_{ij'}-\theta_i)^2\right.\right.$$

$$\left.\left.-2\,(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\right] \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2$$

$$= E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2 \mid y_i.\theta_i,\sigma_i^2;\psi\right\}^2$$

$$+E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2$$
$$+4E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2$$
$$+2E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}$$
$$-4E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\}$$
$$-4E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y'_{ij}-\theta_i)^2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \tag{105}$$

The first term of (105)

$$\begin{aligned}
&= E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2 \\
&= E\left\{(N_i-n_i)\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2 \\
&= (N_i-n_i)^2E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^4+2\sum_{1\le j<k\le n_i}(y_{ij}-\theta_i)^2(y_{ik}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= (N_i-n_i)^2n_i[\mu_4+(n_i-1)\sigma_i^4].
\end{aligned} \tag{106}$$

The second term of (105)

$$\begin{aligned}
&E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2 \\
= {}&E\left\{n_i\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2 \\
= {}&n_i^2[(N_i-n_i)\mu_4+(N_i-n_i)(N_i-n_i-1)\sigma_i^4] \\
= {}&n_i^2(N_i-n_i)[\mu_4+(N_i-n_i-1)\sigma_i^4].
\end{aligned} \tag{107}$$

The third term of (105)

$$E\left[\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\right\}^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right]$$

$$
\begin{aligned}
&= E\left[\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)\right\}^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right] \\
&= E\left[\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2\right\} \mid y_i,\theta_i,\sigma_i^2;\psi\right] \\
&\quad +E\left[\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\sum_{n_i+1\le k\ne p\le N_i}(y_{ik}-\theta_i)(y_{ip}-\theta_i)\right\} \mid y_i,\theta_i,\sigma_i^2;\psi\right] \\
&\quad +E\left[\left\{\sum_{k=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\sum_{1\le j\ne l\le n_i}(y_{ij}-\theta_i)(y_{il}-\theta_i)\right\} \mid y_i,\theta_i,\sigma_i^2;\psi\right] \\
&\quad +E\left[\left\{\sum_{1\le j\ne l\le n_i}(y_{ij}-\theta_i)(y_{il}-\theta_i)\sum_{n_i+1\le k\ne p\le N_i}(y_{ik}-\theta_i)(y_{ip}-\theta_i)\right\} \mid y_i,\theta_i,\sigma_i^2;\psi\right] \\
&= n_i(N_i-n_i)\sigma_i^4, \qquad (108)
\end{aligned}
$$

expectation of the last term in (108) is zero since $y_{ij}-\theta_i$ and $y_{ij'}-\theta_i$ given $(y_i,\theta_i,\sigma_i^2;\psi)$ are independent. The fourth term of (105) becomes

$$
\begin{aligned}
&= 2E\left\{(N_i-n_i)\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2 n_i\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= 2n_i(N_i-n_i)E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= 2n_i(N_i-n_i)E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\}E\left\{\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2 \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= 2n_i^2(N_i-n_i)^2\sigma_i^4. \qquad (109)
\end{aligned}
$$

The fifth term of (105)

$$
\begin{aligned}
&E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= E\left\{(N_i-n_i)\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= (N_i-n_i)E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^3(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&\quad +E\left\{2\sum_{1\le j<k\le n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2(y_{ik}-\theta_i)(y_{ij'}-\theta_i) \mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
&= 0. \qquad (110)
\end{aligned}
$$

The sixth term of (105)

$$
\begin{aligned}
& E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
= \; & n_iE\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
= \; & n_iE\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^3\mid y_i.\theta_i,\sigma_i^2;\psi\right\} \\
& +n_iE\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{n_i+1\le j'\neq k\le N_i}(y_{ik}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
& +n_iE\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)2\sum_{j'<k<l}(y_{ij'}-\theta_i)(y_{ik}-\theta_i)(y_{il}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
= \; & 0. \qquad (111)
\end{aligned}
$$

Using (106)-(111), then the first term of (104) i.e.,

$$
\begin{aligned}
& E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2\mid y_i,\theta_i,\sigma_i^2;\psi\right\}^2 \\
= \; & (N_i-n_i)^2n_i[\mu_4+(n_i-1)\sigma_i^4]+n_i^2(N_i-n_i)[\mu_4+(N_i-n_i-1)\sigma_i^4] \\
& +4n_i(N_i-n_i)\sigma_i^4+2n_i^2(N_i-n_i)^2\sigma_i^4 \\
= \; & n_i(N_i-n_i)[N_i\mu_4+(4N_in_i-4n_i^2+4-N_i)\sigma_i^4]. \qquad (112)
\end{aligned}
$$

The second term of (104)

$$
\begin{aligned}
& E\left\{\sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}(y_{ij}-y_{ij'})^2\mid y_i,\theta_i,\sigma_i^2;\psi\right\} \\
= \; & \sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}E\{(y_{ij}-y_{ij'})^2\mid y_i,\theta_i,\sigma_i^2;\psi\} \\
= \; & \sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}\{E[(y_{ij}-\theta_i)-(y_{ij'}-\theta_i)]^2\mid y_i,\theta_i,\sigma_i^2;\psi\} \\
= \; & \sum_{j=1}^{n_i}\sum_{j'=n_i+1}^{N_i}\Big[E\{(y_{ij}-\theta_i)^2\mid y_i,\theta_i,\sigma_i^2;\psi\}+E\{(y_{ij'}-\theta_i)^2\mid y_i,\theta_i,\sigma_i^2;\psi\} \\
& -2\,E\{(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\}\Big] \\
= \; & 2n_i(N_i-n_i)\sigma_i^2. \qquad (113)
\end{aligned}
$$

the third term in the above is zero since conditional on $(y_i, \theta_i, \sigma_i^2)$ $y_{ij} - \theta_i$ and $y'_{ij} - \theta_i$ are independent. Using (112) and (113), then the first term in (103) is

$$
\begin{aligned}
&= 4\{n_i(N_i - n_i)[N_i\mu_4 + (4N_in_i - 4n_i^2 + 4 - N_i)\sigma_i^4] - (2n_i(N_i - n_i)\sigma_i^2)^2\} \\
&= 4n_i(N_i - n_i)[N_i\mu_4 - (N_i - 4)\sigma_i^4]. \qquad (114)
\end{aligned}
$$

The second term of (103) i.e.,

$$
\begin{aligned}
& Var\left[\sum_{n_i+1\le j\ne j'\le N_i} (y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right] \\
&= Var\left[2(N_i - n_i)\sum_{j=n_i+1}^{N_i} (y_{ij} - \overline{y}'_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right] \\
&= 4(N_i - n_i)^2\left[E\left\{\sum_{j=n_i+1}^{N_i} (y_{ij} - \overline{y}'_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}^2\right. \\
&\quad \left. - \left\{E\sum_{j=n_i+1}^{N_i} [(y_{ij} - \overline{y}'_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi]\right\}^2\right]. \qquad (115)
\end{aligned}
$$

where $\overline{y}'_i = (N_i - n_i)^{-1}\sum_{j=n_i+1}^{N_i} y_{ij}$. The first term of (115) becomes

$$
\begin{aligned}
&= E\left\{\sum_{j=n_i+1}^{N_i} [(y_{ij} - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi] - (N_i - n_i)(\overline{y}'_i - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}^2 \\
&= E\left\{\sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}^2 + (N_i - n_i)^2 E\{(\overline{y}'_i - \theta_i)^4 \mid y_i, \theta_i, \sigma_i^2; \psi\} \\
&\quad -2(N_i - n_i)E\left\{(\overline{y}'_i - \theta_i)^2 \sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}. \qquad (116)
\end{aligned}
$$

The first term of (116) becomes

$$
\begin{aligned}
&= E\left\{[\sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)^4 \mid y_i, \theta_i, \sigma_i^2; \psi]\right. \\
&\quad \left. +[\sum_{n_i+1\le j\ne j'\le N_i} (y_{ij} - \theta_i)^2(y_{ij'} - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi]\right\} \\
&= \sum_{j=n_i+1}^{N_i} E\{(y_{ij} - \theta_i)^4 \mid y_i, \theta_i, \sigma_i^2; \psi\} \\
&\quad + \sum_{n_i+1\le j\ne j'\le N_i} E\{(y_{ij} - \theta_i)^2(y_{ij'} - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\}
\end{aligned}
$$

$$= (N_i - n_i)\mu_4 + (N_i - n_i)(N_i - n_i - 1)\sigma_i^4$$
$$= (N_i - n_i)[\mu_4 + (N_i - n_i - 1)\sigma_i^4]. \tag{117}$$

The second term of (116)

$$= (N_i - n_i)^{-4} E\left[\left\{\sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)\right\}^2 \left\{\sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)\right\}^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right]$$
$$= (N_i - n_i)^{-4} E\left\{\left[\sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)^2 + \sum_{n_i+1 \le j \ne j' \le N_i} (y_{ij} - \theta_i)(y_{ij'} - \theta_i)\right]\right.$$
$$\left.\left[\sum_{k=n_i+1}^{N_i} (y_{ik} - \theta_i)^2 + \sum_{n_i+1 \le s \ne k \le N_i} (y_{ik} - \theta_i)(y_{is} - \theta_i)\right] \mid y_i, \theta_i, \sigma_i^2; \psi\right\}$$
$$= (N_i - n_i)^{-4}\{(N_i - n_i)\mu_4 + (N_i - n_i)(N_i - n_i - 1)\sigma_i^4$$
$$+2(N_i - n_i)(N_i - n_i - 1)\sigma_i^4\}$$
$$= (N_i - n_i)^{-3}\{\mu_4 + 3(N_i - n_i - 1)\sigma_i^4\}. \tag{118}$$

The third term of (116)

$$= (N_i - n_i)^{-2} E\left\{\left[\sum_{j=n_i+1}^{N_i} (y_{ij} - \theta_i)^2 + \sum_{n_i+1 \le j \ne j' \le N_i} (y_{ij} - \theta_i)(y_{ij'} - \theta_i)\right]\right.$$
$$\left.\sum_{j'=n_i+1}^{N_i} (y_{ij} - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}$$
$$= (N_i - n_i)^{-2}\{(N_i - n_i)\mu_4 + (N_i - n_i)(N_i - n_i - 1)\sigma_i^4\}$$
$$= (N_i - n_i)^{-1}\{\mu_4 + (N_i - n_i - 1)\sigma_i^4\}. \tag{119}$$

Using (117), (118) and (119), then the first term of (115) becomes

$$= (N_i - n_i)[\mu_4 + (N_i - n_i - 1)\sigma_i^4] + (N_i - n_i)^{-1}\{\mu_4 + 3(N_i - n_i - 1)\sigma_i^4\}$$
$$-2\{\mu_4 + (N_i - n_i - 1)\sigma_i^4\}$$
$$= (N_i - n_i - 1)^2/(N_i - n_i)\mu_4 + \sigma_i^4(N_i - n_i - 1)[(N_i - n_i) + 3/(N_i - n_i) - 2]. \tag{120}$$

The second term of (115) i.e.,

$$E\left\{\sum_{j=n_i+1}^{N_i} (y_{ij} - \overline{y}_i')^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}$$

$$
\begin{aligned}
&= \sum_{j=n_i+1}^{N_i} \{E(y_{ij}-\theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\} - (N_i - n_i) E\{(\overline{y}_i' - \theta_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\} \\
&= (N_i - n_i)\sigma_i^2 - \sigma_i^2 \\
&= (N_i - n_i - 1)\sigma_i^2. \qquad (121)
\end{aligned}
$$

Now, using (120) and (121) then the second term of (103) i.e.,

$$
\begin{aligned}
& Var\left\{\sum_{n_i+1 \le j \ne j' \le N_i} (y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\} \\
&= 4(N_i - n_i)^2 \left[(N_i - n_i - 1)^2 (N_i - n_i)^{-1}\mu_4 + \sigma_i^4 (N_i - n_i - 1)\right. \\
&\quad \left.(N_i - n_i) + 3/(N_i - n_i) - 2] - (N_i - n_i - 1)^2 \sigma_i^4\right] \\
&= 4(N_i - n_i)(N_i - n_i - 1)\{(N_i - n_i - 1)\mu_4 - (N_i - n_i - 3)\sigma_i^4\}. \qquad (122)
\end{aligned}
$$

The third term of (103) becomes

$$
\begin{aligned}
&= E\left\{\sum_{j=1}^{n_i} \sum_{j'=n_i+1}^{N_i} (y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}, \\
&\quad \left\{\sum_{n_i+1 \le j \ne j' \le N_i} (y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}\Bigg] \\
&\quad - E\left\{\sum_{j=1}^{n_i} \sum_{j'=n_i+1}^{N_i} (y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\} \\
&\quad E\left\{\sum_{n_i+1 \le j \ne j' \le N_i} (y_{ij} - y_{ij'})^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\}. \qquad (123)
\end{aligned}
$$

The first term of (123) becomes

$$
\begin{aligned}
&= E\left\{\sum_{j=1}^{n_i} \sum_{j'=n_i+1}^{N_i} \{(y_{ij} - \theta_i) - (y_{ij'} - \theta_i)\}^2\right. \\
&\quad \left.\sum_{n_i+1 \le j \ne j' \le N_i} \{(y_{ij} - \theta_i) - (y_{ij'} - \theta_i)\}^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\} \\
&= E\left\{\sum_{j=1}^{n_i} \sum_{j'=n_i+1}^{N_i} [(y_{ij} - \theta_i)^2 + (y_{ij'} - \theta_i)^2 - 2(y_{ij} - \theta_i)(y_{ij'} - \theta_i)],\right. \\
&\quad \left.\sum_{n_i+1 \le j \ne j' \le N_i} (y_{ij} - \theta_i)^2 + (y_{ij'} - \theta_i)^2 - 2(y_{ij} - \theta_i)(y_{ij'} - \theta_i) \mid y_i, \theta_i, \sigma_i^2; \psi\right\} \\
&= E\left\{(N_i - n_i)\sum_{j=1}^{n_i} (y_{ij} - \theta_i)^2 + n_i \sum_{j'=n_i+1}^{N_i} (y_{ij'} - \theta_i)^2 - 2\sum_{j=1}^{n_i} (y_{ij} - \theta_i)\right.
\end{aligned}
$$

$$\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i),\ \sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-\theta_i)^2+\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij'}-\theta_i)^2$$

$$-2\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\Bigg\}$$

$$\begin{aligned}
&= E\Bigg\{(N_i-n_i)\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2+n_i\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)^2\\
&\quad -2\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i).2(N_i-n_i-1)\\
&\quad \sum_{j=n_i+1}^{N_i}(y_{ij}-\theta_i)^2-2\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\Bigg\}\\
&= 2(N_i-n_i-1)(N_i-n_i)E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\sum_{j=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\mid y_i,\theta_i,\sigma_i^2;\psi\right\}\\
&\quad +2n_i(N_i-n_i-1)E\left\{\sum_{j=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\sum_{j=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\mid y_i,\theta_i,\sigma_i^2;\psi\right\}\\
&\quad -4(N_i-n_i-1)E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j=n_i+1}^{N_i}(y_{ij}-\theta_i)\sum_{j=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\mid y_i,\theta_i,\sigma_i^2;\psi\right\}\\
&\quad -2(N_i-n_i)E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\}\\
&\quad -2n_iE\left\{\sum_{j'=n_i+1}^{N_i}(y_{ij}-\theta_i)^2\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\}\\
&\quad +4E\left\{\sum_{j=1}^{n_i}(y_{ij}-\theta_i)\sum_{j'=n_i+1}^{N_i}(y_{ij'}-\theta_i)\sum_{n_i+1\le j\ne j'\le N_i}(y_{ij}-\theta_i)(y_{ij'}-\theta_i)\mid y_i,\theta_i,\sigma_i^2;\psi\right\}\\
&= 2(N_i-n_i)(N_i-n_i-1)n_i(N_i-n_i)\sigma_i^4\\
&\quad +2n_i(N_i-n_i-1)\{(N_i-n_i)\mu_4+(N_i-n_i)(N_i-n_i-1)\sigma_i^4\}, \qquad (124)
\end{aligned}$$

since the third to the sixth term in the above are zero. Note that the second term of (123) follows from equation of (113) i.e., $2n_i(N_i-n_i)\sigma_i^2$ and equation of (121) i.e., $(N_i-n_i-1)\sigma_i^2$. Then, using (124), the third term of equation (103) becomes

$$8n_i(N_i-n_i)(N_i-n_i-1)(\mu_4-\sigma_i^4). \qquad (125)$$

Using equation (114), (122) and (125), then the second part in variance of (100)

$$
\begin{aligned}
& Var\left\{\sum_{j=1}^{N_i}(y_{ij}-\bar{y}_i)^2 \mid y_i, \theta_i, \sigma_i^2; \psi\right\} \\
= \; & \frac{1}{4}N_i^{-2}\Big[4(N_i-n_i)n_i(N_i\mu_4-(N_i-4)\sigma_i^4) \\
& +4(N_i-n_i)(N_i-n_i-1)\{(N_i-n_i-1)\mu_4-(N_i-n_i-3)\sigma_i^4\} \\
& +8n_i(N_i-n_i)(N_i-n_i-1)(\mu_4-\sigma_i^4)\Big]. \qquad (126)
\end{aligned}
$$

Note that under normal distribution $\mu_4 = 3\sigma_i^4$, equation (126) becomes

$$
\begin{aligned}
& N_i^{-2}\Big[2(N_i-n_i)n_i(N_i+2)\sigma_i^4 + 2(N_i-n_i)^2(N_i-n_i-1)\sigma_i^4 \\
& + n_i(N_i-n_i)(N_i-n_i-1)4\sigma_i^4\Big] \\
= \; & N_i^{-2}(N_i-n_i)\sigma_i^4\left[2n_i(N_i+2)+2(N_i-n_i-1)(N_i-n_i)(N_i+n_i)\right] \\
= \; & 2N_i^{-1}f_i\sigma_i^4C_i, \qquad (127)
\end{aligned}
$$

where $C_i = n_i(N_i+2)+(N_i+1)(N_i-n_i-1)$. Finally, using (102) and (127) then the variance of the ith population variance estimator is given by

$$
\begin{aligned}
& Var\left\{N_i^{-1}\sum_{j=1}^{N_i}(y_{ij}-\bar{y}_i)^2 \mid y_i\right\} \\
= \; & N_i^{-2}\left[f_i^2(N_i+n_i-1)^2 Var(\sigma_i^2 \mid y_i) + N_i^{-1}f_iC_iE(\sigma_i^4 \mid y_i)\right]. \qquad (128)
\end{aligned}
$$

Table 29: Empirical Bayes Estimate e_i^{EB} of γ_i where $\eta = 10$ and $\xi = 4$

No.	σ_i^2	S_i^2	GL	$e^B(MSE)$	Naive MSE	Boots MSE
1	1.913	1.023	2.388	1.578	0.236	1.861
2	4.006	8.814	3.095	6.183	3.113	1.989
3	6.264	9.868	3.184	6.653	3.531	1.904
4	2.741	3.617	2.621	3.074	0.806	2.210
5	1.264	1.414	2.424	1.807	0.302	2.171
6	2.042	2.527	2.526	2.468	0.537	1.743
7	1.999	1.570	2.441	1.914	0.337	1.726
8	1.659	1.277	2.417	1.754	0.289	1.700
9	3.049	1.691	2.458	2.027	0.380	1.970
10	1.122	0.971	2.384	1.551	0.229	1.845
11	1.050	0.954	2.382	1.538	0.226	1.593
12	2.229	3.991	2.655	3.292	0.917	1.975
13	1.901	2.631	2.533	2.507	0.550	2.082
14	4.619	3.034	2.580	2.844	0.715	1.732
15	1.828	1.440	2.427	1.826	0.308	1.874
16	3.566	3.197	2.606	3.071	0.851	1.888
17	2.211	1.296	2.418	1.762	0.291	1.930
18	2.417	0.785	2.372	1.461	0.208	1.881
19	1.398	1.126	2.400	1.650	0.257	1.822
20	1.162	2.040	2.482	2.186	0.430	1.833
21	1.015	0.952	2.393	1.586	0.243	1.833
22	1.482	1.184	2.403	1.673	0.262	2.043
23	2.816	5.411	2.784	4.125	1.412	2.013
24	3.458	5.506	2.794	4.193	1.460	1.864
25	0.809	1.034	2.389	1.585	0.238	1.885
26	2.636	1.667	2.452	1.986	0.362	1.797
27	1.516	1.459	2.430	1.843	0.314	2.014
28	2.646	1.878	2.467	2.086	0.393	1.798
29	1.774	2.375	2.549	2.716	0.702	1.932
30	2.031	1.978	2.475	0.138	0.411	2.997

Table 30: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of γ_i where $\eta = 5$

	$\lambda = 0.25$				$\lambda = 0.50$			
	S^2	S_i^2	GL	Proposed	S^2	S_i^2	GL	Proposed
AAB	0.1364	0.1118	0.1236	0.0736	0.2729	0.2236	0.2472	0.1459
ASD	0.0279	0.0235	0.0225	0.0105	0.1116	0.0940	0.0898	0.0417
ARB	0.6506	0.3867	0.5824	0.2921	0.6506	0.3867	0.5823	0.2905
ARSD	0.1101	0.0589	0.0881	0.0323	0.2202	0.1178	0.1761	0.0641

	$\lambda = 1.00$				$\lambda = 2.00$			
	S^2	S_i^2	GL	Proposed	S^2	S_i^2	GL	Proposed
AAB	0.5457	0.4473	0.4943	0.2876	1.0915	0.8945	0.9881	0.5649
ASD	0.4464	0.3759	0.3591	0.1642	1.7857	1.5036	1.4353	0.6469
ARB	0.6506	0.3867	0.5821	0.2878	0.6506	0.3867	0.5818	0.2842
ARSD	0.4405	0.2355	0.3520	0.1267	0.8810	0.4710	0.7031	0.2495

	$\lambda = 4.00$			
	S^2	S_i^2	GL	Proposed
AAB	2.1830	1.7890	1.9750	1.1161
ASD	7.1428	6.0143	5.7355	2.5685
ARB	0.6506	0.3867	0.5813	0.2811
ARSD	1.7619	0.9420	1.4040	0.4927

Table 31: The Average Absolute Bias, Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of γ_i where $\eta = 10$

	$\lambda = 0.25$				$\lambda = 0.50$			
	S^2	S_i^2	GL	Proposed	S^2	S_i^2	GL	Proposed
AAB	0.1190	0.1227	0.1039	0.0817	0.2379	0.2453	0.2078	0.1628
ASD	0.0225	0.0333	0.0171	0.0104	0.0899	0.1331	0.0685	0.0412
ARB	0.5468	0.3674	0.4836	0.3173	0.5468	0.3674	0.4835	0.3166
ARSD	0.0915	0.0772	0.0721	0.0338	0.1829	0.1545	0.1441	0.0674

	$\lambda = 1.00$				$\lambda = 2.00$			
	S^2	S_i^2	GL	Proposed	S^2	S_i^2	GL	Proposed
AAB	0.4758	0.4906	0.4155	0.3230	0.9517	0.9812	0.8304	0.6371
ASD	0.3595	0.5322	0.2738	0.1629	1.4381	2.1290	1.0937	0.6385
ARB	0.5468	0.3674	0.4834	0.3154	0.5468	0.3674	0.4830	0.3130
ARSD	0.3659	0.3089	0.2879	0.1340	0.7318	0.6178	0.5748	0.2650

	$\lambda = 4.00$			
	S^2	S_i^2	GL	Proposed
AAB	1.9033	1.9624	1.6585	1.2486
ASD	5.7522	8.5160	4.3651	2.4777
ARB	0.5468	0.3674	0.4823	0.3092
ARSD	1.4635	1.2356	1.1465	0.5189

Table 32: The Average Absolute Bias. Average Square Deviation, Average Relative Bias and Average Relative Square Deviation of different estimates of γ_i where $\eta = 100$

	$\lambda = 0.25$				$\lambda = 0.50$			
	S^2	S_i^2	GL	Proposed	S^2	S_i^2	GL	Proposed
AAB	0.0362	0.0946	0.0343	0.0369	0.0723	0.1893	0.0686	0.0737
ASD	0.0019	0.0125	0.0017	0.0019	0.0074	0.0502	0.0069	0.0077
ARB	0.1472	0.3660	0.1380	0.1491	0.1472	0.3660	0.1380	0.1491
ARSD	0.0077	0.0005	0.0070	0.0079	0.0154	0.0963	0.0139	0.0158

	$\lambda = 1.00$				$\lambda = 2.00$			
	S^2	S_i^2	GL	Proposed	S^2	S_i^2	GL	Proposed
AAB	0.1446	0.3786	0.1374	0.1474	0.2893	0.7571	0.2753	0.2947
ASD	0.0297	0.2006	0.0276	0.0309	0.1188	0.8024	0.1108	0.1236
ARB	0.1472	0.3660	0.1381	0.1490	0.1472	0.3660	0.1384	0.1489
ARSD	0.0307	0.1927	0.0279	0.0316	0.0615	0.3853	0.0561	0.0631

	$\lambda = 4.00$			
	S^2	S_i^2	GL	Proposed
AAB	0.5785	1.5142	0.5518	0.5891
ASD	0.4751	3.2096	0.4454	0.4938
ARB	0.1472	0.3660	0.1386	0.1489
ARSD	0.1230	0.7706	0.1125	0.1261

CHAPTER 6

Beta Binomial in Finite Population Sampling

6.1 Introduction

The methods considered in Chapter 4 and Chapter 5 are valid when the observations are measured in an interval scale. Despite the importance of the analysis of binary data in finite population sampling there is very little emphasis on empirical Bayes estimation which fully specifies a binary model in the estimation procedures. The linear empirical Bayes method of Ghosh and Lahiri (1987) can be used to analyze binary data from a stratified simple random sampling. But due to the use of a robust model, their method does not capture the special feature in binary data. Farrel *et al.* (1992) considered an empirical Bayes method to estimate female labor force participation rates for small areas in the United States. They, however. did not consider empirical Bayes estimation in finite population sampling. Also, a measure of uncertainty of the empirical Bayes estimator which captures all sources of variabilities has not been proposed in the context of finite population proportion estimation.

In section 6.2, we consider the Bayes estimation of finite population proportion from a stratified simple random sample. We consider estimation of prior parameters in section 6.3 and propose an empirical Bayes estimator of finite population proportion. Finally in section 6.4, we propose a measure of uncertainty of empirical Bayes estimator. The proposed measure of uncertainty captures all sources of variation.

6.2 The Bayes Estimation of $\gamma_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$

Let y_{ij} denote the value of a characteristic of interest for the jth unit of the

ith area ($i = 1, \ldots, m; j = 1, \ldots, N_i$). We shall consider the following model:

MODEL 5

(i) Conditional on θ_i , y_{ij}'s are independent with

$y_{ij} \mid \theta_i \sim Bernoulli(\theta_i), \quad (i = 1, \ldots, m; j = 1, \ldots, N_i);$

(ii) $\theta_i \overset{ind}{\sim} Beta(\alpha, \beta), \quad (i = 1, \ldots, m);$

Our objective is to estimate $\gamma_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$ the finite population proportion for the ith stratum ($i = 1, \ldots, m$).

Under the Model 5 and squared error loss, the Bayes estimator of γ_i is given by

$$
\begin{aligned}
e_i^B &= E[\ \gamma_i \mid y_i; \psi] \\
&= N_i^{-1}\left[\sum_{j=1}^{n_i} y_{ij} + \sum_{j=n_i+1}^{N_i} E(y_{ij} \mid y_i; \psi)\right] \\
&= N_i^{-1}\left[\sum_{j=1}^{n_i} y_{ij} + \sum_{j=n_i+1}^{N_i} E\left\{E(y_{ij} \mid y_i, \theta_i; \psi) \mid y_i; \psi\right\}\right] \\
&= N_i^{-1}\left[\sum_{j=1}^{n_i} y_{ij} + \sum_{j=n_i+1}^{N_i} E(\theta_i \mid y_i; \psi)\right]. \qquad (129)
\end{aligned}
$$

Using standard Bayesian calculation (see, e.g., Berger (1985)), it can be shown that

$$
\begin{aligned}
E(\theta_i \mid y_i; \psi) &= \frac{n_i}{n_i + \alpha + \beta)}\overline{y}_i + \frac{\alpha + \beta}{n_i + \alpha + \beta}\frac{\alpha}{\alpha + \beta} \\
&= (1 - w_i)\overline{y}_i + w_i\frac{\alpha}{\alpha + \beta}, \qquad (130)
\end{aligned}
$$

where $\psi = (\alpha, \beta)$, $w_i = \frac{\alpha+\beta}{n_i+\alpha+\beta}$ and $\overline{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij}$. Using (130) in (129) then

$$
\begin{aligned}
e_i^B &= (1 - f_i)\overline{y}_i + f_i((1 - w_i)\overline{y}_i + w_i\frac{\alpha}{\alpha + \beta}) \\
&= (1 - f_i w_i)\overline{y}_i + f_i w_i\frac{\alpha}{\alpha + \beta} \\
&= (1 - B_i)\overline{y}_i + B_i\frac{\alpha}{\alpha + \beta} \qquad (131)
\end{aligned}
$$

Note that $B_i = f_i w_i$ is decreasing function of n_i.

6.3 Empirical Bayes Estimation of $\gamma_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$

To estimate $\psi = (\alpha, \beta)$, we use the method of moment. Note that $\overline{y} = n^{-1}\sum_{i=1}^{m}\sum_{j=1}^{n_i} y_{ij} = n^{-1}\sum_{i=1}^{m} n_i\overline{y}_i$ where $n = \sum_{i=1}^{m} n_i$. Since $E(\overline{y}) = \frac{\alpha}{\alpha+\beta}$ and $MSW = (n-m)^{-1}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2$. Note that $E(MSW) = \sigma^2$ where $\sigma^2 = EVar(y_{ij} \mid \theta_i) = \frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}$. Thus by equating $\frac{\alpha}{\alpha+\beta} = \overline{y}$ and $\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} = MSW$, we get

$$\hat{\alpha} = \frac{\overline{y}\ \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2}{(n-m)\overline{y}(1-\overline{y}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2}$$

$$\hat{\beta} = \frac{(1-\overline{y})\ \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2}{(n-m)\overline{y}(1-\overline{y}) - \sum_{i=1}^{m}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_i)^2} \qquad (132)$$

Now the empirical Bayes estimation of $\gamma_i = N_i^{-1}\sum_{j=1}^{N_i} y_{ij}$ is given by

$$e_i^{EB} = (1 - \hat{B}_i)\overline{y}_i + \hat{B}_i\frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta}} \qquad (133)$$

Note that it is possible to have $\hat{\alpha} < 0$ (which also implies $\hat{\beta} < 0$). In that case, we replace the Beta-Binomial model by the following model:

(i) $y_{ij} \mid \theta_i \overset{ind}{\sim} Bernoulli(\theta_i)$, $(i = 1, \ldots, m; j = 1, \ldots, N_i)$;

(ii) $\theta_i \overset{iid}{\sim} U(0, u)$, $(i = 1, \ldots, m)$,where u is an unknown paremeter and (truncated at u) $0 < u \leq 1$;

Then posterior distribution of θ_i is truncated Beta with parameters $y_i + 1$ and $n_i - y_i + 1$. Then $E(\theta_i \mid y_i; u) = \int_0^u f(\theta_i \mid y_i; u)d\theta_i$ and can be easily found from the Tables for Incomplete Beta Integrals. The unknown parameter u can be estimated by $\hat{u} = 2\overline{y}$.

6.4 Measure of Uncertainty of e_i^{EB}

A measure of variability of e_i^{EB} is given by

$$Var_i^B \ = \ Var_i^B(y_i; \psi) = Var_i^B(\gamma_i \mid y_i; \psi)$$

$$
\begin{aligned}
&= N_i^{-2} Var\left\{ \sum_{j=n_i+1}^{N_i} y_{ij} \mid y_i; \psi \right\} \\
&= N_i^{-2} \left\{ E\left[Var\{ \sum_{j=n_i+1}^{N_i} y_{ij} \mid y_i, \theta_i; \psi\} \mid y_i; \psi \right] \right. \\
&\quad \left. + Var\left[E\{ \sum_{j=n_i+1}^{N_i} y_{ij} \mid y_i, \theta_i; \psi\} \mid y_i; \psi \right] \right\} \\
&= N_i^{-2}[(N_i - n_i)E(\theta_i(1-\theta_i) \mid y_i; \psi) + (N_i - n_i)^2 Var(\theta_i \mid y_i; \psi)] \\
&= N_i^{-1} f_i \frac{y_i + \alpha}{n_i + \alpha + \beta} \frac{n_i + \beta - y_i}{n_i + \alpha + \beta + 1} + f_i^2 \frac{y_i + \alpha}{(n_i + \alpha + \beta)^2} \frac{n_i + \beta - y_i}{n_i + \alpha + \beta + 1} \\
&= \frac{y_i + \alpha}{n_i + \alpha + \beta} \frac{n_i + \beta - y_i}{n_i + \alpha + \beta + 1} [N_i^{-1} f_i + f_i^2 / (n_i + \alpha + \beta)]. \qquad (134)
\end{aligned}
$$

A naive measure of variability of the empirical Bayes estimator e_i^{EB} is obtained as $Var_i^{EB}(y_i; \hat{\psi})$. Note that Var_i^{EB} underestimates the true variability of e_i^{EB} since it does not incorporate the additional variabilities due to estimation of ψ.

Equation (10) of Laird and Louis (1987) can be extended to arrive at the following measure of variability of e_i^{EB}:

$$
Var_i^{EB} = R^{-1} \sum_{r=1}^{R} Var_i^{B}(y_i; \hat{\psi}_r^*) + (R-1)^{-1} \sum_{r=1}^{R} \{e_i^B(y_i, \hat{\psi}_r^*) - \bar{e}_i^B(y_i)\}^2, \qquad (135)
$$

where $\bar{e}_i^B(y_i) = R^{-1} \sum_{r=1}^{R} e_i^B(y_i, \hat{\psi}_r^*)$ and $\hat{\psi}_r^*$ is an estimate of ψ based on the rth bootstrap sample.

BIBLIOGRAPHY

Arora, V. (1994), "Empirical Bayes and Hierarchical Bayes Estimation of Small Area Characteristics," *Ph.D. thesis*, University of Nebraska Lincoln. Dept. of Mathematics and Statistics.

Arora, V., Lahiri, P. and Mukerjee, K. (1997), "Empirical Bayes Estimation of Finite Population Means from Complex Surveys," to appear in *Journal of American Statistical Association.*

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988), "An error-components Model for Prediction of County Crop Areas using Survey and Satellite Data." *Journal of American Statistical Association*, **80**, 28 – 36.

Berger, J.O. (1985), " *Statistical Decision Theory and Bayesian Analysis*" (2nd edn.), Springer-Verlag.

Brackstone, G.J. (1987), "Small area data: Policy Issues and Technical Challenges," *Small Area Statistics: An International Symposium*, Eds. Platek. R., Rao, J.N.K., Särndal, C.E., and Singh, M.P., Wiley, New York, pp. 3 – 20.

Carter, G.M., and Rolph, J.E. (1974), " Empirical Bayes Methods Applied to Estimating Fire Alarm Probabilities", *Journal of the American Statistical Association*, **69**, 880 – 885.

Chaudhuri, A. (1992), " Small Domain Statistics: a review", *Technical Report* ASC/92/2, Indian Statistical Institute, Calcutta.

Cochran, W.G. (1977), " *Sampling Techniques,*" (3rd edn.), New York: John Wiley.

Cox, D.R., and Reid, N. (1987), "Parameter Orthogonality and Approximate Conditional Inference, " *Journal of the Royal Statistical Society*, **B**, **49**, 1 – 39. (with discussion)

Cressie, N. (1992), "REML Estimation in Empirical Bayes Smoothening of Census Undercount, " *Survey Methodology*. **18**, 75 – 94.

Datta, G.S., Ghosh, M. (1993), "Bayesian Estimation of Finite Population Variances with Auxiliary Information," *Sankhyā: The Indian Journal of Statistics* , B, **55**, 156 – 170.

Datta, G.S., Ghosh, M., Huang, E.T., Schultz, L.K., and Tsay, J.H. (1992), "Hierarchical and Empirical Bayes Method for Adjustment of Census Undercount: The 1988 Missouri Dress Rehearsal Data ," *Survey Methodology*, **18**, 95 – 108.

Datta, G.S., Ghosh, M., Nangia, N., and Natarajan, K. (1996), "Estimation of Median Income of four Person Families: a Bayesian Approach," *Tech. Report #429*, Department of Statistics, University of Florida.

Datta, G.S., and Lahiri, P. (1997), "Second Order Approximation to the Mean Squared Error of EBLUP in Small-Area Estimation Problems," unpublished report.

Dick, P. (1995), "Modelling Net Undercoverage in the 1991 Census," *Survey Methodology*, **21**, 45 – 54.

Efron, B., and Morris, C. (1975), "Data Analysis Using Stein's Estimator and its Generalizations, "*Journal of American Statistical Association*, **70**. 311 – 319.

Erdelyi, A. (1956), "*Asymptotic Expansions*," New York: Dover.

Ericksen, E.P. (1974), "A Regression Method for Estimating Population of Local Areas," *Journal of American Statistical Association*, **69**, 867 – 875.

Ericson, W.A. (1969*a*), "Subjective Bayesian Models in Sampling Finite Populations," *Journal of the Royal Statistical Society*, B, **31**, 395 – 233.

Ericson, W.A. (1969*b*), "A Note on the Posterior Mean," *Journal of the Royal Statistical Society*, B, **31**, 332 – 334.

Farrell, P., MacGibbon. B., and Tomberlin, T. (1979), "Empirical Bayes Estimators of Small Area Proportions in Multistage Designs," *Working Paper*

Series, Faculty of Commerce and Administration, Concordia University.

Fay, R. E. (1987), "Application of Multivariate Regression to Small Domain Estimation, " *Small Area Statistics*, Eds., R. Platek, J.N.K. Rao, C.E. Sarndal, and M.P. Singh. Wiley. New Tork, pp. 91 – 102.

Fay, R. E., and Herriot, R.A. (1979), "Estimates of Income for Small Places: an Application of James-Stein Procedure to Census Data, " *Journal of American Statistical Association*, **74**, 269 – 277.

Fay, R. E., Nelson, C.T., and Litow, L. (1993), "Estimation of Median Income for 4-Person Families by State." In *Indirect Estimators in Federal Programs* Statistical Policy Working Paper 21. Statistical Policy Office, Office of Management and Budget, pp. 901 – 917.

Ferguson, W.A. (1973), "A Bayesian Analysis of some Nonparametric Problems. "*The Annals of Statistics*, **1**, 209 – 230.

Ghosh, M., and Lahiri, P. (1987*a*), "Robust Empirical Bayes Estimation of Means from Stratified Samples, "*Journal of American Statistical Association*. **82**, 1153 – 1162.

Ghosh, M., and Lahiri, P. (1987*b*), "Robust Empirical Bayes Estimation of Variance from Stratified Samples, "*Sankhyā*, B, **49**, 78 – 89.

Ghosh, M., and Meeden, G. (1986), "Empirical Bayes Estimation in Finite Population Sampling, "*Journal of American Statistical Association*, **81**. 1058 – 1062.

Ghosh, M., Nangia, N., and Kim, D. (1996), "Estimation of Median Income of Four-Person Families: A Bayesian Time Series Approach, " *Journal of American Statistical Association*, **91**, 1423 – 1431.

Ghosh, M., and Rao, J.N.K. (1994), "Small Area Estimation: An Appraisal". *Statistical Science*, **9**, *No*.1. 55 – 93, (with discussion).

Goldstein, M. (1975), "A Note on Some Bayesian Nonparametric Estimates." *The*

Annals of Statistic, **3**, 736 − 740.

Gonzalez, M. E. (1973), "Use and Evaluation of Synthetic Estimators," *Proceedings of the American Statistical Association, Social Statistics Section*, 33 − 36. Washington D.C.

Hartigan, J.A. (1969), "Linear Bayes methods," *Journal of the Royal Statistical Society, B*, **31**, 446 − 454.

Harville, D.A. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems," *Journal of American Statistical Association*, **72**, 320 − 340.

Henderson, C.R. (1975), "Best Linear Unbiased Estimation and Prediction under a Selection Model," *Biometrics*, **31**, 423 − 447.

Holt, D., Smith, T.M.F., and Tomberlein, T.J. (1979), "A Model-Based Approach to Estimation for Small Subgroups of a Population," *Journal of the American Statistical Association* , **74**, 405 − 410.

Kackar, R.N., and Harville, D.A. (1984), "Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models," *Journal of the American Statistical Association* , **79**, 853 − 862.

Kass, R.E., and Steffey, D. (1989), "Approximate Bayesian Inference in Conditionally Independent Hierarchical Models (Parametric Empirical Bayes Models)," *Journal of the American Statistical Association* , **84**, 717 − 726.

Kleffe, J., and Rao, J.N.K. (1992), "Estimation of Mean Square Error of Empirical Best Unbiased under a Random Error Variance," *Journal of Multivariate Analysis*, **43**, 1 − 15.

Lahiri, P., and Rao, J.N.K. (1995), "Robust Estimation of Mean Squared Error of Small Area Estimators," *Journal of American Statistical Association*, **90**, 758 − 766.

Lahiri, P., and Tiwari, R.C. (1991), "Nonparametrics Bayes and Empirical Bayes

Estimation of Variances from Stratified Samples," *Sankhyā: The Indian Journal of Statistics* , B, **52**, 105 – 118.

Lahiri, P., and Wang, W. (1992), "Estimation of all Employee Links for Small Domains - an Application of Empirical Bayes Procedure," *Proceedings of the Workshop on Statistical Issues in Public Policy Analysis*, Eds.. J.N.K. Rao, and N.G.N. Prasad, **II – 32 – II – 53**.

Laird, N.M., and Louis, T.A. (1987), "Empirical Bayes Confidence Intervals Based on Bootstrap Samples," *Journal of the American Statistical Association* , **82**, 739 – 750.

McCullagh, P. and Zidek, J. (1987), "Regression Methods and Performance Criteria for Small Area Population Estimation," in Eds. Platek, R., Rao, J.N.K., Sārndal, C.E., and Singh, M.P., Wiley, New York, pp. 62 – 74.

Morris, C.N. (1983*a*), "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of American Statistical Association*, **78**, 47 – 59.

Morris, C.N. (1983*b*), "Parametric Empirical Bayes Confidence Intervals," In *Scientific Inference, Data Analysis, and Robustness*, New York; Academic Press, pp. 25 – 50.

Morrison, P. (1971), Demographic Information for Cities: a Manual for Estimating and Projecting Local Population Characteristics. RAND report R-618-HUD.

Nandram, B., and Sedransk, J. (1993), "Empirical Bayes Estimation for the Finite Population Mean on the Current Occasion," *Journal of the American Statistical Association* , **88**, 994 – 1000.

National Research Council (1980), *Panel on Small Area Estimates of Population and Income. Estimating Population and Income of Small Areas*, National Academy Press, Washington, DC.

Prasad, N.G.N, and Rao, J.N.K. (1990), "The Estimation of Mean Squared Errors

of Small Area Estimators, "*Journal of American Statistical Association*. **85**, 163 – 171.

Purcell, N.J. and Kish, L. (1979), "Estimation for Small Domain, "*Biometrics*. **35**, 365 – 384.

Rao, J.N.K. (1986), "Synthetic Estimators, SPREE and Best Model Based Predictors," in *Proceedings of the Conference on Survey Research Methods in Agriculture*, 1 – 16, U.S. Dept. Agriculture, Washington, DC.

Robbins, H.(1955), An Empirical Bayes Approach to Statistics, *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability* (Vol.1), Berkeley: University of California Press 157 – 163.

Royall, R. (1970), "On Finite Population Sampling Theory under Certain Regression Models, "*Biometrika*, **57**, 377 – 387.

Schaible, W.L. (1992), "Use of Small Area Statistics in U.S. Federal Programs." in *Small Area Statistics and Survey Designs* (G. Kalton, J. Kordos and R. Platek, eds.) 1, 95 – 114. Central Statistical Office, Warsaw.

Singh, A.C., Stukel, D.M., and Pfeffermann, D. (1997), "Bayesian Versus Frequentist Measures of Error in Small Area Estimation, " tentatively accepted in *Journal of the Royal Statistical Society*, B.

Skinner, C. (1993), "The Use of Synthetic Estimation Techniques to Produce Small Area Estimates," *Tech. Report*, Dept. of Social Statistics, University of Southampton.

Statistics Canada 1987, *Population Estimation Methods, Canada*, Catalogue 91 – 528*E*. Statistics Canada, Ottawa.

Tierney, L., Kass, R.E., and Kadane, J.B. (1989), "Fully Exponential Laplace Approximations to Expectations and Variances of Nonpositive Functions , "*Journal of American Statistical Association*, **84**, 710 – 716.

Zidek, J.V. (1982), "A Review of Methods for Estimating the Populations of Local

Areas, " *Technical Report* 82 – 4, Univ. British Columbia, Vancouver.

TC2100 PK2

Organization__________________ Bldg./Room__________

U. S. DEPARTMENT OF COMMERCE
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450
IF UNDELIVERABLE RETURN IN TEN DAYS

OFFICIAL BUSINESS

AN EQUAL OPPOI

UNITED STATES POSTAGE
U.S. OFFICIAL MAIL
$ 04.55⁰
MAILED FROM ZIP CODE 22314

UNKNOWN AT THIS P.O. BOX
NAME IS NOT ON BOX APPLICATION

RECEIVED
MAR 2 5 2005
Technology Center 2100